Unit commitment
Progressive Hedging: dealing with binary variables & Chance Constraints

Roger J-B Wets
Mathematics, University of California, Davis

Wednesday, July 4, 2012
Transmission Network

Figure 1. Topology of the IEEE 300 node system
Transmission Network

NE-ISO net
~30,000 BUS

Wednesday, July 4, 2012
In the US is an organization that is responsible for moving electricity over large interstate areas; coordinates, controls and monitors an electricity transmission grid that is larger with much higher voltages than the typical power company's distribution grid.

Is an organization formed at the direction or recommendation of the FERC, in the areas where an ISO is established, it coordinates, controls and monitors the operation of the electrical power system, usually within a single US State, but sometimes encompassing multiple states.

ISO New England Inc. (ISO-NE) is an independent, non-profit RTO, serving Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont. Its Board of Directors and its over 400 employees have no financial interest or ties to any company doing business in the region's wholesale electricity marketplace.
Energy Sources

- nuclear energy
- hydro-power
- thermal plants (coal, oil, shale oil, bio, rubbish, ...)
- gas turbines (natural gas, from "cracking")
- renewables (wind, solar, ..., ocean waves)

different characteristics
Uncertainties

- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators
Uncertainties

- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators
Uncertainties

- WEATHER: demand & supply (especially renewables)
- industrial-commercial environment (demand)
- seasonal, day of the week, time of the day
- contingencies: transmission lines, generators
Market time line

**Day ahead:**
- Post operating reserve requirements
- Prepare and submit DA bids
- Clear DA market using SCUC/SCED
- Post results (DA energy and reserves)
- Re-bidding for RAC
- Post-DA RAC using SCUC

**Operating day:**
- Intraday RAC using SCUC
- Clear RT market using SCED (every 5 min)
- Prepare and submit RT bids
- Post results (RT energy and reserves)

Short history of ISO-management techniques

- **RT**: deterministic optimization with LMP (dual variables associated with demand(s) constraints).
- **SCUC/SCED**: Lagrangian relaxation with conservative reliability constraints
- **SCUC/SCED**: deterministic MIP with conservative RUT
- **ARPA E (project)**: “take into account uncertainty”
A collection of sto-programs

- DA-SCUC/SCED unit commitment  \( \text{binaries} \)
- DA-RAC rebidding assessment bidding  \( \text{(binaries)} \)
- DA-RUT - reliability commitments (spinning, N-1)
- RT - 3 min (real time adjustments) LMP’s
- SCED2 - 3 or 4 hours schedule to foresee ramp ups/down, etc.

\( DA = \text{day ahead} \)
Team composition:

- SCUC/SCED model designers + optim. implementation: UCD (Woodruff + , Wets), Sandia National Labs (Watson, Silvia, Siirola, Ross + Sandia Livermore).

- Uncertainty description: UCD (Wets), Iowa State (Ryan, Tesfatsion, Alipantis + )

- Software prototype: Alstom (Kwok Cheung +)

- ISO-mentor: NE-ISO (Eugene Litvinov + ...)

- Market modifications: Iowa State (Tesfatsion, Alipantis + all)
Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

SCUC enforces limited number of transmission constraints on the commitment solution.

Day-Ahead Market

Each hourly SCED performs SFT, which tests all contingencies in a list and for violations, imposes appropriate constraints in SCED and resolves it.

Abstract Unit Commitment

Minimize \( \sum_{k \in K} \sum_{j \in J} c^p_j(k) + c^u_j(k) + c^d_j(k) \) with

- \( K \) time periods
- \( J \) generating units

\[ \text{power output} \quad \sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K \]

\[ \text{demand} \]

\[ \text{max power output} \quad \sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K \]

\[ \text{spinning reserve} \]

\[ p_j(k), \bar{p}_j(k) \in \Pi, \quad \forall j \in J, \forall k \in K \]

\[ \Pi \text{ region of feasible production, all generating units, all time periods.} \]

The specific nature of \( \Pi \) is model-dependent.
Abstract Unit Commitment

Minimize \[ \sum_{k \in K} \sum_{j \in J} c_j^P(k) + c_j^u(k) + c_j^d(k) \]

with

- Production cost
- Startup cost
- Shutdown cost

K time periods \( j \) generating units

\[ \sum_{j \in J} p_j(k) = D(k), \quad \forall k \in K \]

Demand

Max power output

\[ \sum_{j \in J} \bar{p}_j(k) \geq D(k) + R(k), \quad \forall k \in K \]

Spinning reserve

\[ p_j(k), \bar{p}_j(k) \in \Pi, \quad \forall j \in J, \quad \forall k \in K \]

\( \Pi \) region of feasible production, all generating units, all time periods.

The specific nature of \( \Pi \) is model-dependent.

"Stochastic Version"
between a rock and a hard place

CPLEX-MIP: can handle a few scenarios

PH: not designed for binary variables
Progressive Hedging Algorithm

0. \( w^0_{\xi} \) such that \( \mathbb{E}\{w^0_{\xi}\} = 0, \nu = 0 \). Pick \( \rho > 0 \)

1. for all \( \xi \):

\[
(x^{1,\nu}_{\xi}, x^{2,\nu}_{\xi}) \in \arg \min f(\xi; x^1, x^2) + \langle w^\nu_{\xi}, x^1 \rangle + \frac{\rho}{2} \| x^1 - \bar{x}^{1,\nu-1} \|^2
\]

\[ x^1 \in C^1 \subset \mathbb{R}^{n_1}, x^2 \in C^2(\xi, x^1) \subset \mathbb{R}^{n_2} \]

2. \( \bar{x}^{1,\nu} = \mathbb{E}\{x^{1,\nu}_{\xi}\} \). Stop if \( \| x^{1,\nu}_{\xi} - \bar{x}^{1,\nu} \| = 0 \) (approx.)

otherwise \( w^{\nu+1}_{\xi} = w^\nu_{\xi} + \rho \left[ x^{1,\nu}_{\xi} - \bar{x}^{1,\nu} \right] \), return to 1. with \( \nu = \nu + 1 \)

Implementation: bundling, \( \rho \rightarrow \rho_s \), ...

Watson & Woodruff (Hart, Siirola, ...)

Chile: Sistemas Complejos de Ingenieria (L.F. Solari, ...)

& Centro de Modelamiento Matematico

Carl Laird (Texas A&M), Ryan Sarah (Iowa), ...
min\langle c, x\rangle + \sum_{\xi \in \Xi} p_\xi \langle q_\xi, y_\xi \rangle \text{ such that } \\
x \in C_1, \ y_\xi \in C_2(\xi, x) \ \forall \ \xi \in \Xi \\
binary (integer) variables: some \ x’s, some \ y_\xi’s.
**PH: binary variables**

\[
\min \langle c, x \rangle + \sum_{\xi \in \Xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle \\
\text{such that} \\
x \in C_1, \ y_{\xi} \in C_2(\xi, x) \ \forall \xi \in \Xi
\]

binary (integer) variables: some \(x\)'s, some \(y_{\xi}\)'s.

Choice of \(\rho \rightarrow \rho_j\) depending on \(c_j, |x_j|, ...\)

Variable Fixing, in particular binaries, \(x_j(s) = \text{constant (}k\text{ iterations)}\)

Variable Slamming: aggressive variable fixing \(x_j(s) \approx \text{constant (} & c_j x_j(s)\text{)}\)

“Sufficient” variable convergence \(\sim\) for small values of \(c_j x_j(s)\)

Termination criterion: variable slamming when \(x_j^\nu(\xi) - x_j^{\nu+1}(\xi)\) small

Detecting cycling behavior: (simple) hashing scheme
PH: binary variables

\[
\min \langle c, x \rangle + \sum_{\xi \in \Xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle \text{ such that }
\begin{align*}
x &\in C_1, \quad y_{\xi} \in C_2(\xi, x) \quad \forall \xi \in \Xi \\
\text{binary (integer) variables: some } x\text{'s, some } y_{\xi}\text{'s.}
\end{align*}
\]

Choice of \( \rho \to \rho_j \) depending on \( c_j, |x_j|, \ldots \)

Variable Fixing, in particular binaries, \( x_j(s) = \text{constant (} k \text{ iterations)} \)

Variable Slamming: aggressive variable fixing \( x_j(s) \approx \text{constant (} \& c_j x_j(s) \))

“Sufficient” variable convergence \( \sim \) for small values of \( c_j x_j(s) \)

Termination criterion: variable slamming when \( x_j^\nu(\xi) - x_j^{\nu+1}(\xi) \) small

Detecting cycling behavior: (simple) hashing scheme

**Enough variables fixed \( \Rightarrow \) clean up with CPLEX-MIP**
Large Scale Chance-Constraints
min $\langle c, x \rangle + \sum_{\xi \in \Xi} d_{\xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle$

such that $(x, y_{\xi}) \in C_{\xi}$, $\forall \xi \in [S : d_{\xi} = 1]$

$\sum_{\xi \in S} p_{\xi} d_{\xi} \geq 1 - \alpha$, $d_{\xi} \in \{0, 1\}, \xi \in S$

**Aircraft sustainability problem:** min $\langle c, x \rangle$ such that

$A_{\xi} x \geq b_{\xi} d_{\xi}$, $\forall \xi \in S$, $x \in \mathbb{R}_+$

$d_{\xi} \in \{0, 1\}, \forall \xi \in S$, $\sum_{\xi \in S} d_{\xi} \geq (1 - \alpha)|S|$

$x$: inventory policy, resource levels, time-index variables ($> 10^6$)

$A_{\xi}, b_{\xi}$ discrete event simulation of operation sustainability

$|\Xi| \approx 5 \cdot 10^6$, $\alpha \sim 0.04$. 

Wednesday, July 4, 2012
PH & G.Chance Constraints

Relaxation:
\[
\begin{align*}
\min & \langle c, x \rangle + \sum_{\xi \in \Xi} d_{\xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle - \lambda \left( \sum_{\xi \in S} p_{\xi} d_{\xi} - (1 - \alpha) \right) \\
\text{such that } & (x, y_{\xi}) \in C_{\xi}, \quad \forall \xi \in [S : d_{\xi} = 1] \\
& d_{\xi} \in \{0, 1\}, \xi \in S
\end{align*}
\]

for all $\lambda \geq 0$ yields a lower bound for the generalized C.C. problem
Relaxation:
\[
\min \langle c, x \rangle + \sum_{\xi \in \Xi} d_{\xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle - \lambda \left( \sum_{\xi \in S} p_{\xi} d_{\xi} - (1 - \alpha) \right)
\]
such that \((x, y_{\xi}) \in C_{\xi}, \quad \forall \xi \in [S : d_{\xi} = 1] \)
\[
d_{\xi} \in \{0, 1\}, \xi \in S
\]
for all \(\lambda \geq 0\) yields a lower bound for the generalized C.C. problem

Ignoring coupling C.C., for \(\xi \in S\): let
\[
(\bar{x}_{\xi}, y_{\xi}) \in \arg\min \langle c, x \rangle + \langle q_{\xi}, y \rangle
\]
such that \((x, y) \in C_{\xi}\)
Relaxation:
\[
\min \langle c, x \rangle + \sum_{\xi \in \Xi} d_\xi p_\xi \langle q_\xi, y_\xi \rangle - \lambda \left( \sum_{\xi \in S} p_\xi d_\xi - (1 - \alpha) \right)
\]
such that \((x, y_\xi) \in C_\xi, \ \forall \xi \in [S : d_\xi = 1]\)
\[d_\xi \in \{0, 1\}, \xi \in S\]

for all \(\lambda \geq 0\) yields a lower bound for the generalized C.C. problem

Ignoring coupling C.C., for \(\xi \in S\): let
\[(\bar{x}_\xi, y_\xi) \in \arg\min \langle c, x \rangle + \langle q_\xi, y \rangle\]
such that \((x, y) \in C_\xi\)

”Decomposed” calculation of \(d_\xi\):
\[
\min \langle c, \bar{x}_\xi \rangle - \lambda d_\xi
\]
such that \((\bar{x}_\xi, y_\xi) \in C_\xi, \ d_\xi \in \{0, 1\}\)
PH & G.Chance Constraints

Relaxation:
\[ \min \langle c, x \rangle + \sum_{\xi \in \Xi} d_{\xi} p_{\xi} \langle q_{\xi}, y_{\xi} \rangle - \lambda \left( \sum_{\xi \in S} p_{\xi} d_{\xi} - (1 - \alpha) \right) \]
\[ \text{such that } (x, y_{\xi}) \in C_{\xi}, \quad \forall \xi \in [S : d_{\xi} = 1] \]
\[ d_{\xi} \in \{0, 1\}, \xi \in S \]

for all \( \lambda \geq 0 \) yields a lower bound for the generalized C.C. problem

Ignoring coupling C.C., for \( \xi \in S \): let
\[ (x_{\xi}, y_{\xi}) \in \text{argmin} \langle c, x \rangle + \langle q_{\xi}, y \rangle \]
\[ \text{such that } (x, y) \in C_{\xi} \]

"Decomposed" calculation of \( d_{\xi} \):
\[ \min \langle c, x_{\xi} \rangle - \lambda d_{\xi} \]
\[ \text{such that } (x_{\xi}, y_{\xi}) \in C_{\xi}, \quad d_{\xi} \in \{0, 1\} \]

Solution: let
\[ d_{\xi} = 1 \text{ when } \langle c, x_{\xi} \rangle + \langle q_{\xi}, y_{\xi} \rangle \leq \lambda \]
otherwise \( d_{\xi} = 0 \)
assuming \( \lambda \) given

0. \( w_0^\xi \) such that \( \mathbb{E}\{w_0^\xi\} = 0 \), \( \nu = 0 \). Pick \( \rho > 0 \)

1. for all \( \xi \in S \):

\[
(x_\xi^\nu, y_\xi^\nu) \in \arg \min \langle c, x \rangle + \langle q, y \rangle + \langle w_\xi^\nu, x \rangle + \frac{\rho}{2} \left| x^1 - \bar{x}^{1,\nu-1} \right|^2, \quad (x, y) \in C_\xi
\]

2. if \( \langle c, x_\xi^\nu \rangle + \langle q, y_\xi^\nu \rangle \leq \lambda \), set \( d_\xi = 1 \) otherwise \( d_\xi = 0 \)

3. \( \bar{x}^\nu = \left( \sum_{\xi \in S} p_\xi d_\xi x_\xi^\nu \right) / \left( \sum_{\xi \in S} p_\xi d_\xi \right) \) and

\[
w_{\xi}^{\nu+1} = w_\xi^\nu + \rho \left[ x_\xi^{1,\nu} - \bar{x}^{1,\nu} \right]
\]

4. if \( \left( \frac{1}{\sum_{\xi \in S} p_\xi d_\xi} \sum_{\xi \in S} p_\xi d_\xi \right) \sum_{\xi \in S} p_\xi d_\xi | x_\xi^\nu - \bar{x}^\nu | > \varepsilon \); return to 1. with \( \nu = \nu + 1 \)

otherwise Stop
Biasing the $d_\xi$ variables

Strategy: PH $\rightarrow x_\xi^*, \lambda_{\text{max}} \searrow \lambda^*$:

augmentation function (minimization of discontinuous functions):

limit function: $\tau = \langle c, x_\xi \rangle + \langle q_\xi, y_\xi \rangle$

$$\psi(\tau, \lambda, \lambda_{\text{max}}) = \begin{cases} 1 & \text{when } 0 \leq \tau \leq \lambda, \\ 0 & \text{for } \lambda < \tau \leq \lambda_{\text{max}} \end{cases}$$

mollifiers: Beta densities on $[0, \Delta]$, $\Delta \searrow 0$ as PH converges (gap)

$$\varphi(\Delta; z) = \begin{cases} \frac{1/\Delta}{B(7,1.75)} (z/\Delta)^6 (1 - z/\Delta)^{0.75} & \text{when } z \in [0, \Delta] \\ 0 & \text{elsewhere} \end{cases}$$

Beta function $B(a, b) = \int_0^1 z^{a-1} (1 - z)^{b-1} \, dz$
Augmentation function

\[ m(\Delta, \lambda_{\text{max}}; z, \lambda) = \int_{0}^{\Delta} \psi(z - s, \lambda, \lambda_{\text{max}}) \varphi(\Delta; s) \, ds, \quad z \in [0, \lambda_{\text{max}}] \]
Augmentation function

\[ m(\Delta, \lambda_{\text{max}}; z, \lambda) = \int_0^\Delta \psi(z - s, \lambda, \lambda_{\text{max}}) \varphi(\Delta; s) \, ds, \quad z \in [0, \lambda_{\text{max}}] \]
0. $w^0_\xi$ such that $\mathbb{E}\{w^0_\xi\} = 0$, $\nu = 0$. Pick $\rho > 0$, $\epsilon > 0$, $\Delta = 1$

1. for all $\xi \in S$: ($\nu = 0$ ignore the proximal term, fix $\lambda_{\text{max}}$ an upper bound on cost fcn)

$$\left(x^\nu_\xi, y^\nu_\xi\right) \in \arg\min \left\langle c, x \right\rangle + \left\langle q, y \right\rangle + \left\langle w^\nu_\xi, x \right\rangle + \frac{\rho}{2} \left| x^1 - \bar{x}^{1,\nu - 1} \right|^2, \quad (x, y) \in C_\xi$$

2. $(\lambda, d) = \arg\min_{\lambda \leq \lambda_{\text{max}}} \lambda$ such that for all $\xi \in S$,

$$d_\xi = m\left(\Delta, \lambda_{\text{max}}; \left\langle c, x^{\nu - 1}_\xi \right\rangle + \left\langle q_\xi, y_\xi \right\rangle, \lambda\right) \quad \& \quad \sum_{\xi \in S} p_\xi d_\xi \geq 1 - \alpha$$

3. $\bar{x}^\nu = \left(\sum_{\xi \in S} p_\xi d_\xi x^\nu_\xi\right) / \left(\sum_{\xi \in S} p_\xi d_\xi\right)$ and

$$w^{\nu + 1}_\xi = w^\nu_\xi + \rho \left[ x^{1,\nu}_\xi - \bar{x}^{1,\nu} \right]$$

4. if $\left(1 / \sum_{\xi \in S} p_\xi d_\xi\right) \sum_{\xi \in S} p_\xi d_\xi \mid x^\nu_\xi - \bar{x}^\nu \mid > \epsilon$; return to 1. with $\nu = \nu + 1$

otherwise Stop
Unit Commitment
Modeling Load
HOW to model stochastic processes?
Example: load in Connecticut zone of ISO-NE

• Data from ISO-NE: For each of 8 load zones,
  – Date
  – Hour: 1-24
  – Temperature: Dry bulb in deg. F
  – Dew Point: Dew point temperature in deg. F
  – Demand

• CT accounts for 27.7% of electricity sales
• Consider data since 2006 after major market changes implemented by ISO-NE in 2005

6/12/2012
Load Modeling Process
Example: load in Connecticut zone of ISO-NE

- Exploratory data analysis to determine major influences
- Data segmentation
- Multiple linear regression (MLR) within data segments to determine relationships
- Also experimented with:
  - Time series transfer functions within data segments
  - Semi-parametric time series approach
  - Multiple linear regression on whole data set with dummy variables for hour, day-type and month

(excruciating data analysis)
CT load exploratory process

Hourly Load in A Week

Seasonal Trend of Hourly Load in Days

Weekly Pattern by Month, CT 2011
CT load vs. weather variables

Load vs. temperature (TMP) and dew point temperature (DPT)

Load vs. TMP (DPT=19)
Load vs. TMP (DPT=52)

Load vs. DPT in Jan
Load vs. DPT in Aug

Load vs. TMP in Mar
Load vs. TMP in Jul

6/12/2012

ARPA-e Project Review
Tentative CT Load Model

• Data segmented by season, day-type and hour
• For each segment, fit MLR model on 2006-2010 training data set

\[ L(k) = \beta_0 + \beta_1 \text{TMP}(k) + \beta_2 \text{TMP}(k)^2 + \beta_3 \text{TMP}(k)^3 + \beta_4 \text{DPT}(k) + \beta_5 DPT(k)^2 + \beta_6 \text{DPT}(k)^3 \\
+ \beta_7 \text{TMP}(k-1) + \beta_8 \text{TMP}(k-2) + \cdots + \beta_{14} \text{TMP}(k-8) + \beta_{15} \text{TMP}(k-24) + \beta_{16} \text{TMP}(k-168) \\
+ \beta_{17} \text{DPT}(k-1) + \beta_{18} \text{DPT}(k-2) + \cdots + \beta_{24} \text{DPT}(k-8) + \beta_{25} \text{DPT}(k-24) + \beta_{26} \text{DPT}(k-168) \\
+ \varepsilon(k) \]

• Average relative error on training set \( \sim 3\% \)
• Mean absolute percent error (MAPE) on 2011 test data \( \sim 6\% \)
• To do: replace \( \text{TMP}(k-m), \text{DPT}(k-m) \) terms with \( L(k-m) \);
  consider interactions \( \text{TMP}(k) \times \text{DPT}(k) \)

⇒ Generating scenarios