Abridged Nested Decomposition for Multistage Stochastic Linear Programming

John R. Birge

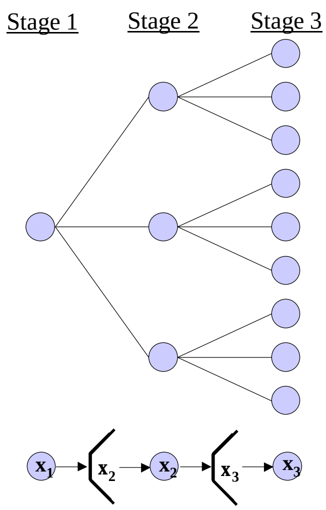
University of Michigan
Industrial and Operations Engineering

May 1999

Outline

- Multistage Stochastic Linear Program
- Nested Decomposition
- Model for Colombian Generation Plan
- Abridged Nested Decomposition
- Computational Results
- Summary

Multistage Stochastic Linear Program



$$\begin{array}{lll} \min & c_1 x_1 + Q_2(x_1) & \\ s.t. & W_1 x_1 = & h_1 \\ & x_1 \geq & 0 \end{array}$$

$$Q_t\left(x_{t-1,a(k)}\right) = \sum_{x_{t,k} \in \Xi_t} prob(x_{t,k}) \mathrm{Q}_{t,k}\left(x_{t-1,a(k)}, x_{t,k}\right)$$

$$\mathrm{Q}_{t,k}\left(x_{t-1,a(k)}, x_{t,k}\right) = \begin{array}{lll} \min & c_t(x_{t,k}) x_{t,k} + Q_{t+1}(x_{t,k}) & \\ s.t. & W_t x_{t,k} = & h_t(x_{t,k}) - T_{t-1}(x_{t,k}) x_{t-1,a(k)} \\ & x_{t,k} \geq & 0 \end{array}$$

- $Q_{N+1}(x_N) = 0$, for all x_N,
- $Q_{t,k}(x_{t-1,a(k)})$ is a piecewise linear, convex function of $x_{t-1,a(k)}$

Feasibility Cuts

- **Consider any stage N problem $Q_{N,k}(x^*_{N-1,a(k)}, x_{N,k})$**

$$\begin{aligned} Q_{N,k}\left(x^*_{N-1,a(k)}, x_{N,k}\right) &= \min\left\{c_N(x_{N,k}) x_{N,k} \mid W_N x_{N,k} = h_N(x_{N,k}) - T_{N-1}(x_{N,k}) x^*_{N-1,a(k)}, x_{N,k} \geq 0\right\} \\ &= \max\left\{p_{N,k}\left(h_N(x_{N,k}) - T_{N-1}(x_{N,k}) x^*_{N-1,a(k)}\right) \mid p_{N,k} W_N \leq c_N(x_{N,k})\right\} \end{aligned}$$

- **Infeasibility of $Q_{N,k}(x^*_{N-1,a(k)}, x_{N,k})$ implies that there exists a direction of unboundness (p^*) such that**

$$p^*\left(h_N(x_{N,k}) - T_{N-1}(x_{N,k}) x^*_{N-1,a(k)}\right) > 0$$

- **Adding the following feasibility cut to stage N-1 subproblem ($a(k)$) makes $x^*_{N-1,a(k)}$ infeasible**

$$p^* T_{N-1}(x_{N,k}) x_{N-1,a(k)} \geq p^* h_N(x_{N,k})$$

Optimality Cuts

- **Consider the Stage *N-1* expected recourse function from any stage *N-1* node *k***

$$Q_N\left(x^*_{N-1,a(k)}\right) = \sum_{x_{N,k}\in\Xi_{N,k}} prob\left(x_{N,k}\right) Q_N\left(x^*_{N-1,a(k)}, x_{N,k}\right)$$
$$= \sum_{x_{N,k}\in\Xi_{N,k}} prob\left(x_{N,k}\right) * \left(\min\left\{c_N\left(x_{N,k}\right)x_{N,k} \mid W_N x_{N,k} = h_N\left(x_{N,k}\right) - T_{N-1}\left(x_{N,k}\right)x^*_{N-1,a(k)}, x_{N,k} \geq 0\right\}\right)$$
$$= \sum_{x_{N,k}\in\Xi_{N,k}} prob\left(x_{N,k}\right) * \left(\max\left\{p_{N,k}\left(h_N\left(x_{N,k}\right) - T_{N-1}\left(x_{N,k}\right)x^*_{N-1,a(k)}\right) \mid p_{N,k} W_N \leq c_N\left(x_{N,k}\right)\right\}\right)$$

- **Let (p*) denote optimal dual solutions. Then,**

$$Q_N\left(x_{N-1,a(k)}\right) \geq \sum_{x_{N,k}\in\Xi_{N,k}} prob\left(x_{N,k}\right) * \left(p^*_{N,k}\left(h_N\left(x_{N,k}\right) - T_{N-1}\left(x_{N,k}\right)x_{N-1,a(k)}\right)\right)$$

Nested Decomposition

- **In each subproblem, replace expected recourse function $Q_{t,k}(x_{t-1,a(k)})$ with unrestricted variable $q_{t,k}$**
  - **Forward Pass:**
    - **Starting at the root node and proceeding forward through the scenario tree, solve each node subproblem**

$$\begin{aligned}
\hat{Q}_{t,k}\left(x_{t-1,a(k)}, x_{t,k}\right) = \quad & \min \quad c_t\left(x_{t,k}\right)x_{t,k} + q_{t,k} \\
& s.t. \quad W_t x_{t,k} = h_t\left(x_{t,k}\right) - T_{t-1}\left(x_{t,k}\right)x_{t-1,a(k)} \\
& \qquad E_{t,k} x_{t,k} + q_{t,k} \geq e_{t,k} \quad (\textit{optimality cuts}) \\
& \qquad D_{t,k} x_{t,k} \geq d_{t,k} \quad (\textit{feasibility cuts}) \\
& \qquad x_{t,k} \geq 0
\end{aligned}$$

    - **Add feasibility cuts as infeasibilities arise**
  - **Backward Pass**
    - **Starting in top node of Stage t = N-1, use optimal dual values in descendant Stage t+1 nodes to construct new optimality cut. Repeat for all nodes in Stage t, resolve all Stage t nodes, then t ⟶ t-1.**
  - **Convergence achieved when** $q_1 = Q_2(x_1)$

Pereira-Pinto Method

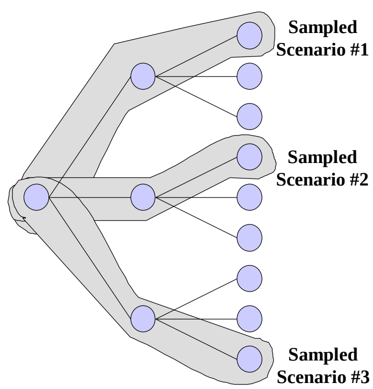
- Incorporates sampling into the general framework of the Nested Decomposition algorithm
- Assumptions:
  - relatively complete recourse
    - no feasibility cuts needed
  - serial independence
    - an optimality cut generated for any Stage t node is valid for all Stage t nodes
- Successfully applied to multistage stochastic water resource problems

Pereira-Pinto Method

1. Randomly select *H N*-Stage scenarios
2. Starting at the root, a forward pass is made through the sampled portion of the scenario tree (solving ND subproblems)
3. A statistical estimate of the first stage objective value $\bar{z}$ is calculated using the total objective value obtained in each sampled scenario

the algorithm terminates if current first stage objective value $c_1x_1 + q_1$ is within a specified confidence interval of $\bar{z}$

4. Starting in sampled node of Stage t = N-1, solve all Stage t+1 descendant nodes and construct new optimality cut. Repeat for all sampled nodes in Stage t, then repeat for t = t - 1



Pereira-Pinto Method

- Advantages
  - significantly reduces computation by eliminating a large portion of the scenario tree in the forward pass
- Disadvantages
  - requires a complete backward pass on all sampled scenarios
    - not well designed for bushier scenario trees

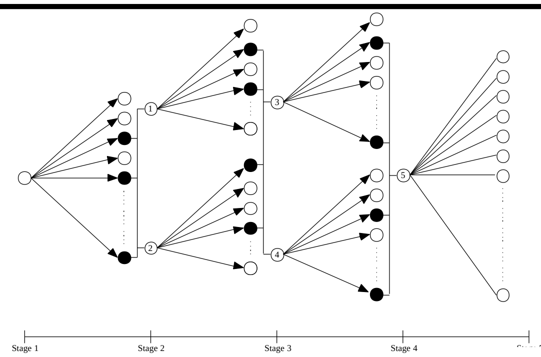
Abridged Nested Decomposition

- Also incorporates sampling into the general framework of Nested Decomposition
- Also assumes relatively complete recourse and serial independence
- Samples both the subproblems to solve and the solutions to continue from in the forward pass

Abridged Nested Decomposition

Forward Pass

1. **Solve root node subproblem**
2. **Sample Stage 2 subproblems and solve selected subset**
3. **Sample Stage 2 subproblem solutions and branch in Stage 3 only from selected subset (i.e., nodes 1 and 2)**

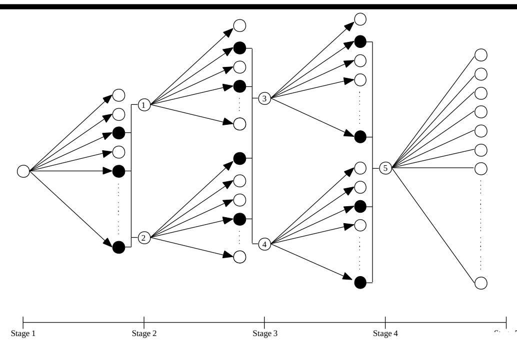


4. **For each selected Stage t-1 subproblem solution, sample Stage t subproblems and solve selected subset**
5. **Sample Stage t subproblem solutions and branch in Stage t+1 only from selected subset**

Abridged Nested Decomposition

Backward Pass

1. **Starting in first branching node of Stage t = N-1, solve all Stage t+1 descendant nodes and construct new optimality cut for all stage t subproblems. Repeat for all sampled nodes in Stage t, then repeat for t = t - 1**



Convergence Test

1. **Randomly select *H N*-Stage scenarios. For each sampled scenario, solve subproblems from root to leaf to obtain total objective value for scenario**
2. **Calculate statistical estimate of the first stage objective value $\bar{z}$**
   - **algorithm terminates if current first stage objective value $c_1x_1 + q_1$ is within a specified confidence interval of $\bar{z}$; else, a new forward pass begins**

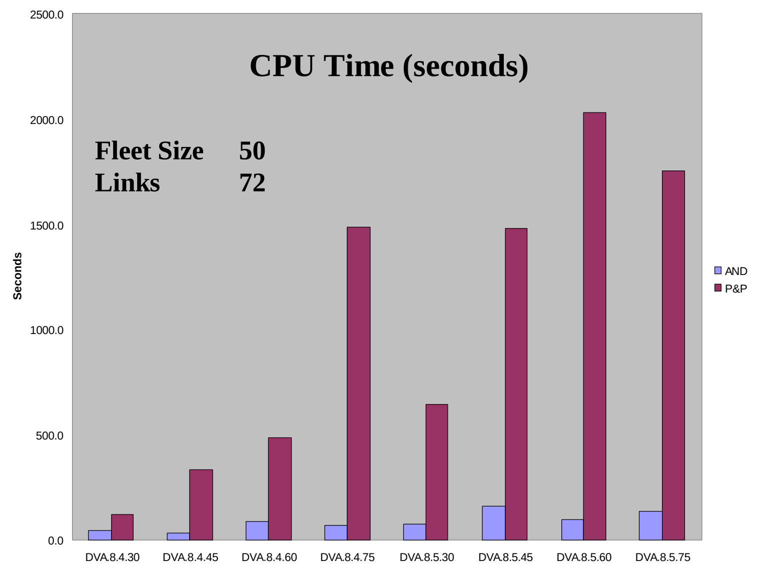
Computational Results

- **Implementation of Pereira & Pinto Method and Abridged Nested Decomposition**
  - **written in C, run on Sun SPARC 20 workstation**
  - **uses CPLEX to solve subproblems**
- **Pereira & Pinto Method**
  - **uses a sample size of 30 for each problem**
- **Abridged Nested Decomposition**
  - **number of Stage t subproblems solved from each Stage t-1 branching value: 15**
  - **initial number of Stage t branching values: 2**
    - **number of Stage t branching values increases with each failed convergence test**
- **Both methods terminate when first stage objective value is within one standard deviation of statistical estimate**

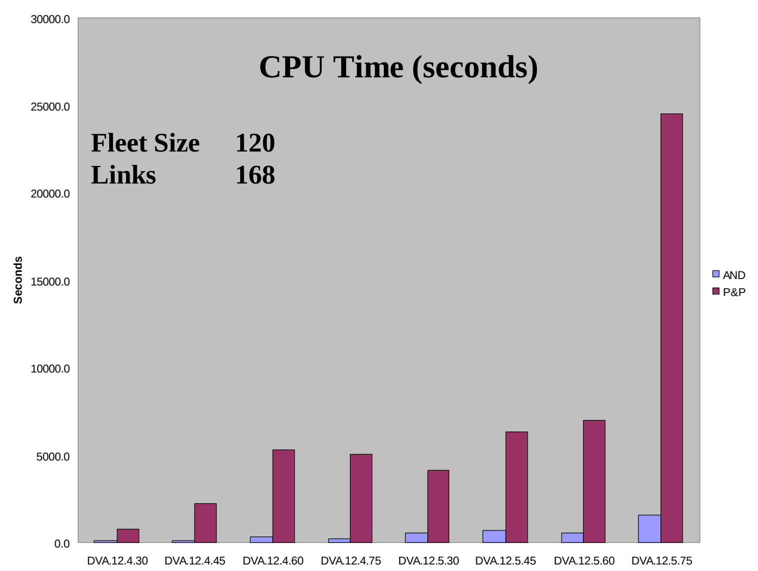
Computational Results

- **Test Problems**
  - **Dynamic Vehicle Allocation (DVA) problems of various sizes**
    - **set of homogeneous vehicles move full loads between set of sites**
    - **vehicles can move empty or loaded, remain stationary**
    - **demand to move load between two sites is stochastic**
  - **DVA.*x.y.z***
    - ***x* number of sites (8, 12, 16)**
    - ***y* number of stages (4, 5)**
    - ***z* number of distinct realizations per stage (30, 45, 60, 75)**
  - **largest problem has > 30 million scenarios**

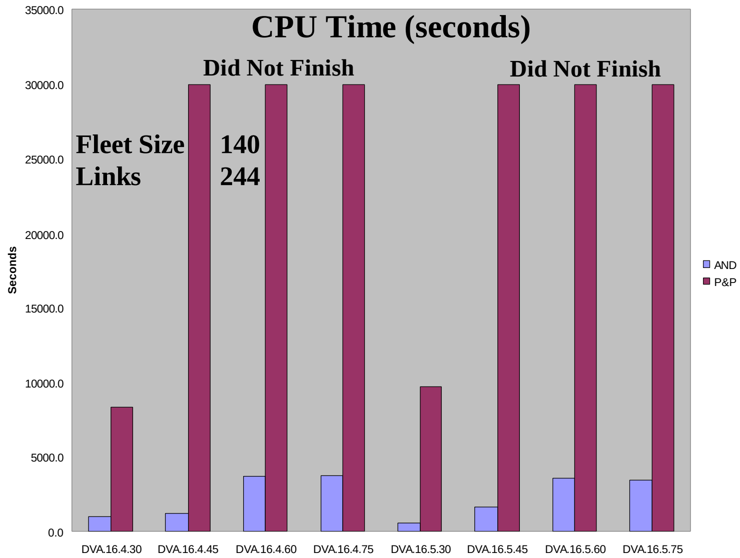
Computational Results (DVA.8)



Computational Results (DVA.12)



Computational Results (DVA.16)



Model for Colombia Power

- **Basic model in AMPL**
- **Solvable in CPLEX**
- **Assume serial correlation within a larger stage (may extend over several months)**
- **Single branch or multiple branches**
- **Abridged NDUM able to solve to optimality**

Using Abridged NDUM

- **Compile in C with proper links to CPLEX directories**
- **Create linked object code in AND**
- **Use files for each period in filei.mps**
- **Use Sdata for stochastics**
- **Use Tdata for linking matrices**

Additional Programs

- **C code to create data files based on databases**
- **AMPL code to create mps files for ND**
- **CPLEX models in fullmodel.mod and full.dat**
- **Tried 3, 10 and 23 years of data**
- **Also, used for single scenario for capacity evaluation**

Basic Model

- **Objective:**
  - minimize total generation costs
- subject to the operating constraints:
  - meet load constraints
  - thermal capacity constraints
  - hydro maximum/minimum flow constraints
  - export/import capacity constraints
  - minimum/maximum reservoir level
  - penalty on alert levels and minimum levels

Penalty Calculations

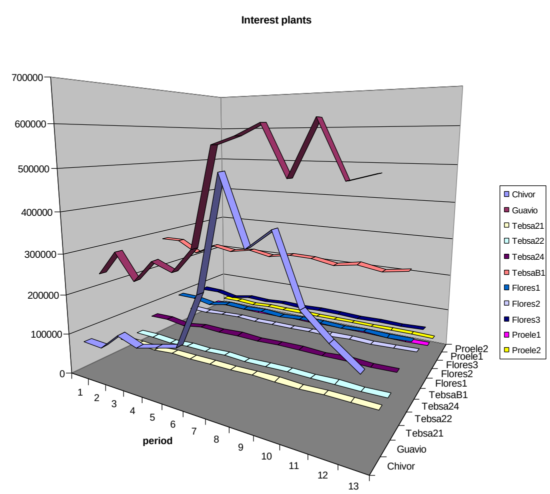
- **Rationing Costs (minimum level violated)**

SEGMENT	$/Mwh	range (% of unsatisfied demand)
1	308022	0 - 1.5
2	558435	1.5 - 5
3	979351.9	5 - 90
4	1939307.3	90 - 100

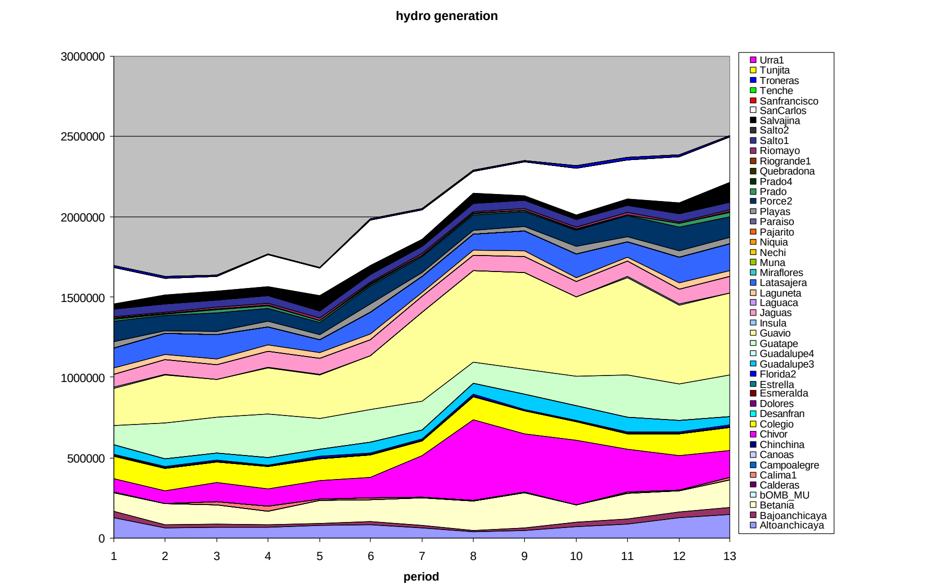
Alert Level Penalties

- **Calculation:**
  - **Start at penalty (per MWhr) equal to highest thermal unit cost (dual on meet-load constraint)**
  - **Interpolate to first block outage cost**
- **Procedure:**
  - **Assign penalty to meet-load dual**
  - **Re-solve problem and iterate until solution converges (3-5 iterations in general)**

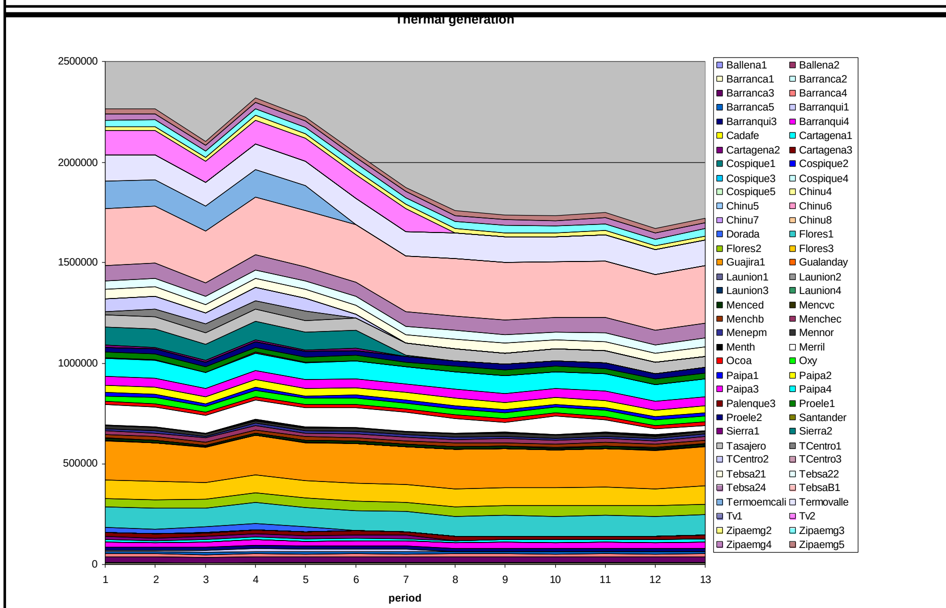
Example Results (selected plants)



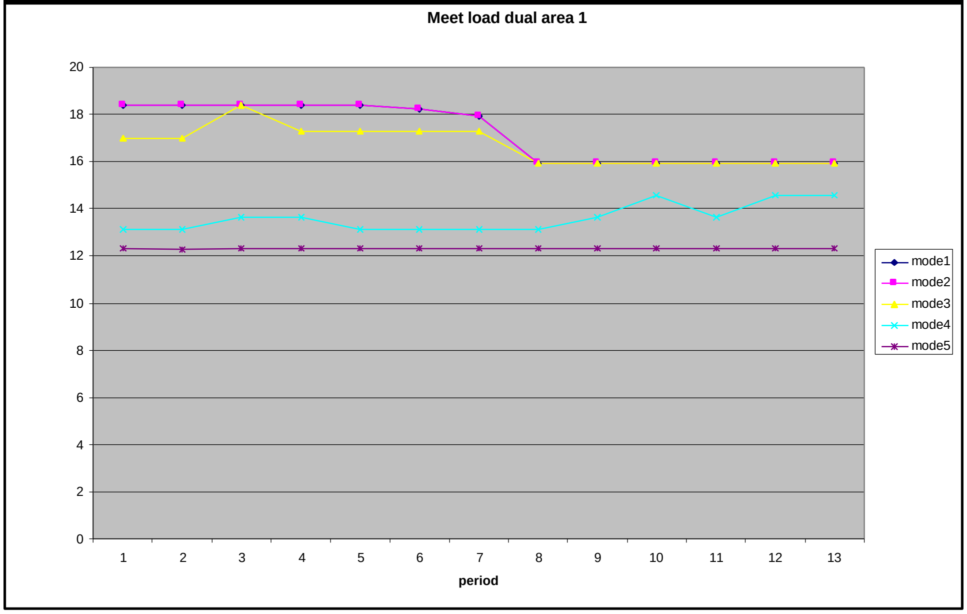
Example: Hydro Generation



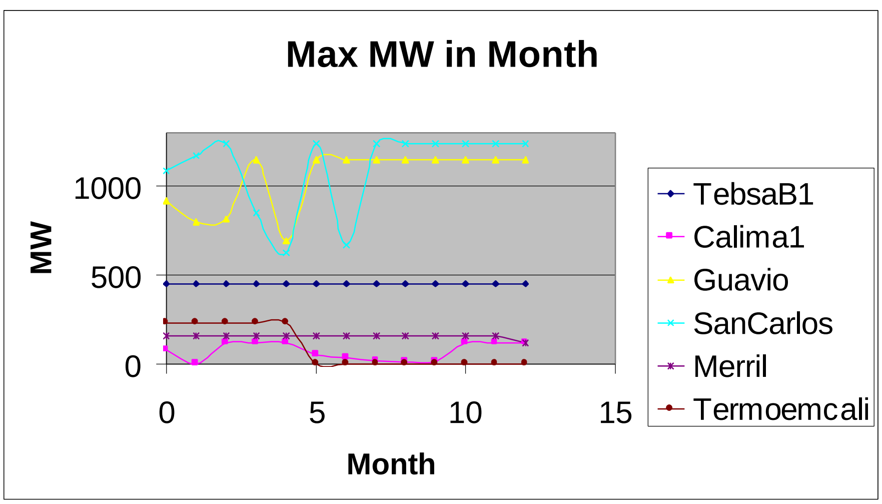
Example: Thermal Generation



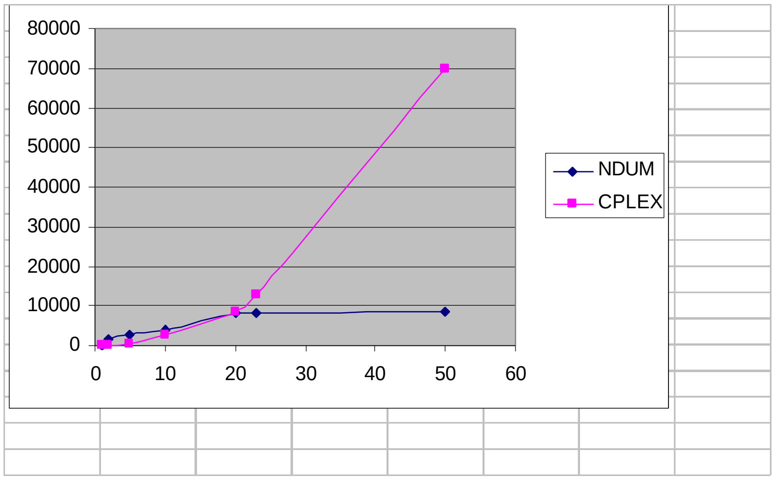
Example: Dual Prices



Example: Max MW (selected plants)



NDUM and CPLEX v. No. of Scenarios



Summary

- **Both Pereira-Pinto method and Abridged Nested Decomposition method incorporate sampling into the framework of Nested Decomposition method**
  - **both assume relatively complete recourse and serial independence among effective stages**
- **Pereira-Pinto method is well-designed for long, narrow scenario trees**
- **Abridged Nested Decomposition method well-designed for bushier scenario tree**