MIT ESD

Massachusetts Institute of Technology
Engineering Systems Division

ESD Working Paper Series

Heterogeneous Unit Clustering for Efficient Operational Flexibility Modeling for Strategic Models

Bryan S. Palmintier
Engineering Systems Division
Massachusetts Institute of Technology
Cambridge, MA 02139 USA
E-mail: b_p@mit.edu

Mort D. Webster
Engineering Systems Division
Massachusetts Institute of Technology
Cambridge, MA 02139 USA
Phone: 617-253-3901
E-mail: mort@mit.edu

ESD-WP-2013-04 January 2013 esd.mit.edu/wps

Heterogeneous Unit Clustering for Efficient Operational Flexibility Modeling for Strategic Models

Bryan S. Palmintier and Mort D. Webster

***Abstract*— The increasing penetration of wind generation has led to significant improvements in unit commitment models. However, long-term capacity planning methods have not been similarly modified to address the challenges of a system with a large fraction of generation from variable sources. Designing future capacity mixes with adequate flexibility requires an embedded approximation of the unit commitment problem to capture operating constraints. Here we propose a method, based on clustering units, for a simplified unit commitment model with dramatic improvements in solution time that enable its use as a submodel within a capacity expansion framework. Heterogeneous clustering speeds computation by aggregating similar but non-identical units thereby replacing large numbers of binary commitment variables with fewer integers that still capture individual unit decisions and constraints. We demonstrate the trade-off between accuracy and run-time for different levels of aggregation. A numeric example using an ERCOT-based 205-unit system illustrates that careful aggregation introduces errors of 0.05-0.9% across several metrics while providing several orders of magnitude faster solution times (400x) compared to traditional binary formulations and further aggregation increases errors slightly (~2x) with further speedup (2000x). We also compare other simplifications that can provide an additional order of magnitude speed-up for some problems.**

***Index Terms*—Integer programming, Power generation scheduling, Power system modeling, Unit commitment, Flexibility, Capacity Expansion.**

NOMENCLATURE

A. Indices

g, G	Generating unit, set of units
$\hat{g}$, $\hat{G}$	Generation cluster, set of clusters
t, τ, T	Time period
ρ	Reserve category {1,2,3}

B. Variables

C^{total}	Total system cost [$]
$C^{var}_{g,t}$	Variable costs [$]
$C^{start}_{g,t}$	Startup costs [$]
$P_{g,t}$	Power output [MWh]
F_g	Fuel usage [mmbtu/MWh]
$U_{g,t}$	0/1 unit commitment state
$S_{g,t}$	0/1 startup indicator
$D_{g,t}$	0/1 shutdown indicator
$R^{1,up}_{g,t}$	Primary (regulation) up reserves
$R^{1,down}_{g,t}$	Primary (regulation) down reserves
$R^{2,up}_{g,t}$	Secondary (load follow) up reserves
$R^{2,down}_{g,t}$	Secondary (load follow) down reserves
$R^{3}_{g,t}$	Tertiary reserves (quick start)

C. Parameters

c^{fuel}_g	Fuel costs [$/mmbtu]
$c^{varO\&M}_g$	Variable O&M costs [$/mmbtu]
f^{start}_g	Fuel usage at startup [mmbtu/MWh]
l_t	Load [MWh]
p^{min}_g	Minimum power output [MWh]
p^{max}_g	Maximum power output [MWh]
Δp^{down}_g	Maximum down-ramp rate [MWh/hr]
Δp^{up}_g	Maximum up-ramp rate [MWh/hr]
$r^{1,up}$	Primary reserve requirement up [MWh]
$r^{1,down}$	Primary reserve requirement down [MWh]
$r^{2,up}$	Secondary reserve requirement up [MWh]
$r^{2,down}$	Secondary reserve req. down [MWh]
r^{outage}	Contingency reserve requirement [MWh]
$r^{replace}$	Tertiary reserve requirement [MWh]
x^{nosync}	Fraction of secondary reserves from offline

This work was supported in part by the U.S. National Science Foundation under Grants 1128147 and 0835414.

B. S. Palmintier (e-mail: b_p@mit.edu) and M. D. Webster (e-mail: mort@mit.edu) are with the Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

a_g^{ρ}	Unit capability by reserve class [MWh]
$a_g^{quickstart}$	0/1 quick start ability
$m_g^{up/down}$	Minimum up/downtime [hrs]

I. Introduction

Growth in variable renewables and other advanced power system technologies has prompted a need for researchers to capture operational flexibility in a range of models. Operational flexibility requires a balance between 1) requirements due to uncertainty (e.g. forecast errors and outages) plus dynamics (e.g. demand and wind ramps) and 2) limitations, typically from thermal generator technical constraints (e.g. minimum output levels, startup/shutdown limits, maximum ramping, etc.). In flexibility studies, these dynamics and constraints are typically captured using unit commitment (UC) models [1–18].

Since some of the early pioneering work in UC models [1-2], there have been significant contributions in reformulating unit commitment models to appropriately represent variable generation and its impacts on reserve requirements and operations within an existing capacity mix [3-14]. Much of this work has been to develop improved algorithms for stochastic unit commitment, inclusion of transmission and security constraints, and the use of these models to develop optimal reserve allocation rules and to economically value the additional reserve requirements from renewables. In addition, there have been initial attempts to consider long-term flexibility needs within capacity planning models [15-17].

However, the unit commitment (UC) problem is by itself computationally intensive to solve because of the combination of the large number of discrete (binary) on/off decisions – one for each generator for each time period – and the number and complexity of the technical constraints. To include the full UC problem within a long-term planning model would make the solution of these models infeasible because the UC subproblem has to be solved for many alternative capacity mixes considered, often over a long timeframe (e.g., one year). What is needed within capacity expansion models is a simplified approximation of the UC problem that captures its large scale features.

Because the goal here is to study future potential systems that do not yet exist, the level of detail for required for current systems, such as transmission and security constraints and generator-specific characteristics, is not warranted. The lack of detailed unit-specific data enables the aggregation or clustering of similar generation units, which transforms the large number of binary commitment variables to far fewer integer variables, and thereby drastically reducing problem size and corresponding run times [15], [18], [19].

The concept of aggregating *identical* units is not new. As early as 1966, pioneering studies in computer based unit commitment, grouped identical generators together to illustrate simple solution techniques with limited computer hardware [20]. More recently, examples of combining identical units has also appeared in the literature. For example, Gollmer, et al. [19] also use grouped integer commitment for identical thermal plants and Garcia-Gonzalez, et al. [21] use a grouped integer on/off state when modeling banks of identical hydro turbines for optimal combined bidding with wind. Likely other implementations using such homogeneous clustering remain unpublished, since the computational advantages of binary aggregation to integers is well recognized in the operations research community [22]. For example in his dissertation, Cerisola describes homogeneous aggregation into “generalized” units with integer commitment variables [23], yet this formulation is not described in related journal articles [24]. When clustering identical, co-located units, clustering can provide identical solutions in faster times.

The concept of *heterogeneous* clustering extends this aggregation such that similar, but not identical, units are clustered together and assigned an integer commitment state. Conceptually, this approach is similar to that of Sen and Kothari [25], who also group units. However, their treatment assumes a binary commitment state for the entire group: all on or all off. This is computationally helpful, but is much less flexible than an integer formulation that allows some of the generators within a group to run while others are off. The all or none approach also prevents properly computing startup costs, minimum output levels, and reserve capability.

Recent work on heterogeneous clustering has demonstrated efficient unit-commitment-based computations over long time horizons (e.g. full year as 8760 sequential hours) as part of price estimation [18] and planning studies [15]. However, in both efforts, heterogeneous clusters are simply used to make the study tractable, without considering different clustering approaches or comparing the results to a full binary formulation.

A key contribution of this paper is to explore the trade-offs among accuracy, run-time and level of aggregation used in heterogeneous clustering. To do so, we introduce a set of performance metrics applicable to a wide range of decision objectives. We also compare the performance of other simplifying long-term UC assumptions with and without clustering.

In addition, this paper presents a streamlined implementation for clustered minimum up and down time, which uses only one integer variable per cluster. As described in [18] and [23], these dynamic, inter-period constraints require careful consideration since within the same cluster, some units may startup or shutdown while others continue to run. In the past, these constraints have been converted back to binary [23], not described fully [19], or are not relevant because non-thermal units are aggregated [21]. Langrene et al. describe heterogeneous clustered dynamic constraints in detail, but their minimum up/down formulation requires multiple integer variables per cluster per time period [18]. These separate integers explicitly represent startup/shutdown generator states: running but stoppable, must keep running, stopped but able to start, and must stay stopped. As described below, our formulation uses sums of existing continuous startup and shutdown variables to maintain only a single integer unit commitment variable per cluster, thereby further reducing the problem size.

In Section II, we formulate a standard binary UC model that we use to compare against our clustered formulation. We present the clustered formulation of the UC model in Section III. Section IV describes alternative speedup strategies used in the literature to which we compare our approach. The experimental setup and error metrics are defined in Section V. Section VI presents the results for a 205 generator unit test system. Section VII provides a concluding discussion.

II. Traditional Unit Commitment

A. Core model

The generic unit commitment problem finds the minimum cost combination of generator commitment and power output to meet demand over time. Here we linearize the standard basic formulation [26], [27], for a thermal-only system. The resulting optimization problem is a large mixed-integer linear program (MILP) that can then be solved by powerful commercial solvers as is done by a growing number of power system operators [24]. For clarity, we use uppercase for variables, bold upper case for sets, and lowercase for parameters and set elements.

1) The Objective Function minimizes total operations costs:

$$C^{total} = \min \sum_{g\in \boldsymbol{G}} \sum_{t\in \boldsymbol{T}} \left(C_{g,t}^{var} + C_{g,t}^{start}\right) \tag{1}$$

computed as the sum of variable costs, $C_{g,t}^{var}$, and startup costs, $C_{g,t}^{start}$, for all units, g, and time periods, t.

2) The Variable Costs, $C_{g,t}^{var}$ include fuel costs, c_g^{fuel}, and variable operations and maintenance (O&M) costs, $c_g^{varO\&M}$, as a function of the instantaneous power output, $P_{g,t}$:

$$\begin{aligned} &C_{g,t}^{var} = F_g P_{g,t} c_g^{fuel} + P_{g,t} c_g^{varO\&M} \\ &\text{with} \quad P_{g,t} \geq 0, \end{aligned} \tag{2}$$

3) The Startup Costs, $C_{g,t}^{start}$, assume a constant fuel use per startup, f_g^{start}:

$$C_{g,t}^{start} = S_{g,t} f_g^{start} c_g^{fuel} \tag{3}$$

Startup events, $S_{g,t}$, are computed using the state equation:

$$U_{g,t} = U_{g,t-1} + S_{g,t} - D_{g,t} \tag{4}$$

$$\text{with} \quad U_{g,t}, S_{g,t}, D_{g,t} \in \{0,1\} \tag{5a}$$

Here $U_{g,t}$ represents the commitment (on/off) state of each unit, $S_{g,t}$ represents startup events, and $D_{g,t}$ represents a unit shut down; all are 0/1 variables.

We note that (3) deviates from startup formulations that distinguish warm and cold startup costs, e.g. [26]. This constant startup cost simplification is commonly used for this class of long-term unit commitment problem [14], [17], [28].

4) A Piecewise Linear Fuel Usage function, $F_g(P_{g,t})$, captures the non-linear relation between fuel usage and power output. We represent this using a unit-specific convex piecewise linear approximation with segments, $\boldsymbol{B_g}$:

$$F_{g,t}\left(P_{g,t}\right) \geq h_{g,b} P_{g,t} + U_{g,t} f_{g,b}^{P=0} \qquad \forall s \in \boldsymbol{B_g} \tag{6}$$

For each piecewise linear segment, b, the slope, $h_{g,b}$, represents the incremental heat rate and the intercept, $f_{g,b}^{P=0}$, indicates the projected fuel use if hypothetically running at zero power. Since fuel has a positive cost, the optimizer will minimize fuel usage forcing the inequality in equation (6) to equality for the highest piecewise segment. When a unit is not running, the commitment variable, $U_{g,t}$, brings fuel use to zero.

5) The System Balance Constraint ensures that the sum of instantaneous power, $P_{g,t}$, equals total load, l_t, at all times:

$$\sum_{g\in \boldsymbol{G}} P_{g,t} = l_t \qquad \forall t \in \boldsymbol{T} \tag{6}$$

6) Unit Minimum and Maximum Output Constraints use the binary commitment variable to imply that each generating unit is either off and outputting zero power ($U_{g,t} = 0$), or on and running within its operating limits, p_g^{min} and p_g^{max} ($U_{g,t} = 1$):

$$U_{g,t} p_g^{min} \leq P_{g,t} \leq U_{g,t} p_g^{max} \tag{7a}$$

B. Additional Constraints

A more realistic model includes additional cost components and generator and system reliability imposed technical constraints [27]. We focus on the most common extensions:

1) Ramping Limits capture limitations on how fast thermal units can adjust their output power:

$$P_{g,t-1} - P_{g,t} \leq \Delta p_g^{downmax} \tag{8a}$$

$$P_{g,t+1} - P_{g,t} \leq \Delta p_g^{upmax} \tag{9a}$$

where the Δp's are the ramp limits up or down.

2) For Minimum Up and Down Times we adopt the most computationally efficient formulation from [29], [30] using, m_g^{up} and m_g^{down} for minimum up and down times:

$$U_{g,t} \geq \sum_{\tau = t - m_g^{up}}^{t} S_{g,\tau} \tag{10}$$

$$1 - U_{g,t} \geq \sum_{\tau = t - m_g^{down}}^{t} D_{g,\tau} \tag{11a}$$

3) Operating Reserves: Because power generated on the grid must match demand instantaneously, a number of operating reserves are maintained by allowing room between generator output levels and corresponding limits to provide on-line capacity able to quickly increase (or decrease) and compensate for generation or transmission outages, forecast errors, etc.:

a) *Primary reserves,* which operate on a timescale of a few seconds to compensate for rapid stochastic changes:

$$\sum_{g\in \boldsymbol{G}} R_{g,t}^{1,up} \geq r^{1,up} l_t \tag{12}$$

$$\sum_{g\in \boldsymbol{G}} R_{g,t}^{1,down} \geq r^{1,down} l_t \tag{13}$$

where $R_{g,t}^{1,up}$ and $R_{g,t}^{1,down}$ are the quantities of primary reserves supplied by unit g in time period t. The totals of which must exceed the exogenously determined system-level frequency reserve requirements, $r^{1,up}$ and $r^{1,down}$.

b) *Secondary reserves* operate on the few minute timescale for both contingencies (spinning reserves) and load following. We allow a fraction of the reserve up supply, x^{nosync}, to be supplied by non-synchronized resources such as offline quick starting units or demand response.:

$$\sum_{g\in \boldsymbol{G}} R_{g,t}^{2,up} \geq \left(r^{2,up} l_t + r^{outage}\right)\left(1 - x^{nosync}\right) \tag{14}$$

$$\sum_{g\in\boldsymbol{G}} R_{g,t}^{2,down} \geq\ r^{2,down} l_t \tag{15}$$

The $R_{g,t}$'s are the quantity of on-line secondary reserves supplied by each unit. $r^{2,up}$ and $r^{2,down}$ are the system load following requirements, a function of load/wind forecast error. r^{outage} is the additional reserve required for contingencies, typically set to the largest unit or transmission tie capacity.

c) Tertiary or quick start reserves are off-line but ready to run units that can be brought on-line quickly when needed:

$$\sum_{g\in\boldsymbol{G}} \left(R_{g,t}^{3} + R_{g,t}^{2,up}\right) \geq\ r^{2,up} D_t + r^{outage} + r^{replace} \tag{16}$$

The left-hand side includes both tertiary and secondary up reserves to both capture the fraction of the secondary reserve allowed by (15) from off-line units, and to enable tertiary reserves to be met by on-line units when appropriate.

d) *Unit reserve capabilities* are dictated by the units ability to provide each type of reserve, a_g^{ρ}:

$$R_{g,t}^{\rho,[up/down]} \leq a_g^{\rho,[up/down]} p_g^{max} \quad \forall \rho \in \{1,2\} \tag{17a}$$

For tertiary reserves, quick start capable units can only be drawn from the pool of non-active units:

$$R_{g,t}^{3} \leq (1 - U_{g,t}) a_g^{quickstart} p_g^{max} \tag{18a}$$

where $a_g^{quickstart}$ represents the fraction of the unit capacity, p_g^{max}, that can be deployed fast enough.

e) Updated unit output constraints capture the need for a unit to run below maximum for upward and above minimum for downward reserves. This replaces (7a) with the pair:

$$\begin{aligned} P_{g,t} &\geq U_{g,t} p_g^{min} + R_{g,t}^{1,down} + R_{g,t}^{2,down} \\ U_{g,t} p_g^{max} &\geq P_{g,t} + R_{g,t}^{1,up} + R_{g,t}^{2,up} \end{aligned} \tag{7b}$$

III. Clustered Unit Commitment

A. The Concept of Clustering

As described in the introduction, for problems with simplified or ignored transmission constraints, it is possible to combine similar generating units into clusters. As seen in Fig. 1, this replaces the large set of *binary* commitment decisions, one for each unit, with a smaller set of *integer* commitment states, one for each cluster. All of the other variables – such as power output level, reserves contribution, etc. – and constraints are then aggregated for the entire cluster. Within the cluster, however, the integer commitment variable still captures individual unit level relations.

Computationally, the integer variables provide structure that both reduces the dimensionality of and guides the search through the combinatorial commitment state space by eliminating identical or very similar permutations of binary commitment decisions. The number of possible discrete combinations of commitment variables with the traditional formulation scales exponentially as 2^N with the number of units N, while clustering scales as the product of the cluster sizes: $\prod n_{\hat{g}}$. For example, a system with 100 units clustered into three groups of sizes {10, 70, 20} would reduce the

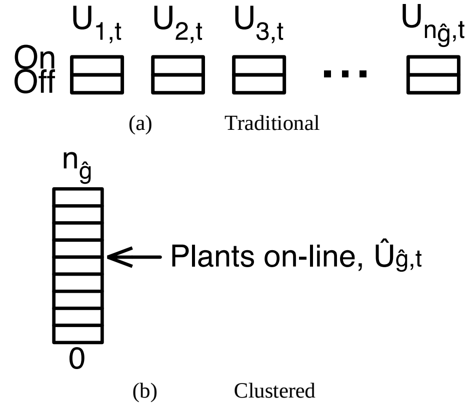


Fig. 1. Conceptual comparison between traditional and clustered unit commitment for a single type of unit in a single time period. In the traditional formulation (a), each unit has a separate binary commitment variable, $U_{g,t}$. With clustering (b), the entire cluster of $n_{\hat{g}}$ units has only a single integer commitment variable, $\widehat{U}_{\hat{g},t}$.

number of discrete combinations in each time period from $\sim 10^{30}$ to $\sim 10^{4}$.[1] In addition, clustering reduces the number of continuous equations and variables since all relations now apply over the smaller number of clusters rather than the full set of individual units.

B. Clustering Formulation

Mathematically, little of the traditional formulation changes with clustering. The key exceptions are replacing the individual unit index, g, with the cluster identifier, $\hat{g}$, and using a corresponding integer commitment variable, $\widehat{U}_{\hat{g}}$:

$$\widehat{U}_{\hat{g}} \in \{0,1,\dots,n_{\hat{g}}\} \tag{5b}$$

1) Relations With No Change Needed. Beyond this substitution no further changes are required for the objective (1), variable costs (2), startup costs (3), commitment state (4), system balance (6), unit output constraints (7b), minimum up time (10), and system reserve requirements (12) – (16).

2) Ramping Limits require the most extensive changes since hour-to-hour output for the entire cluster must account for unit start up, $S_{\hat{g},t}$, and shut down, $D_{\hat{g},t}$. The ramp rates for on-line generators also scale by the number of plants actually on-line, $\hat{U}_{\hat{g},t}$. These modify (8a) & (9a) to:

$$\begin{aligned} &P_{\hat{g},t-1} - P_{\hat{g},t} \\ &\quad \leq \widehat{U}_{\hat{g},t} \Delta p_g^{downmax} + \max(p_g^{min}, \Delta p_g^{downmax}) D_{\hat{g},t+1} \end{aligned} \tag{8b}$$

$$\begin{aligned} &P_{\hat{g},t+1} - P_{\hat{g},t} \\ &\quad \leq \widehat{U}_{\hat{g},t} \Delta p_g^{upmax} + \max(p_g^{min}, \Delta p_g^{upmax}) S_{\hat{g},t+1} \end{aligned} \tag{9b}$$

3) The Minimum Down Time requires finding the number of units currently off as the difference between $n_{\hat{g}}$ (as opposed to one) and the current commitment state, $\widehat{U}_{\hat{g},t}$:

$$n_{\hat{g}} - \widehat{U}_{\hat{g},t} \geq \sum_{\tau = t - m_{\hat{g}}^{mindown}}^{t} D_{\hat{g},\tau} \tag{11b}$$

4) Reserve capabilities change similarly to:

[1] Modern MILP solvers use sophisticated branch-and-cut algorithms to explore only a tiny fraction of this combinatorial space. Still, the speedup with reduced dimensionality can be dramatic.

$$\mathrm{R}_{\hat{g},t}^{\rho} \leq \widehat{U}_{\hat{g},t} a_g^{\rho} p_g^{max} \quad \forall \rho \in \{1, down; 2, down; 1, up;\ 2, up\} \tag{17b}$$

$$R_{\hat{g},t}^{tertiary} \leq (n_{\hat{g}} - \widehat{U}_{\hat{g},t}) a_g^{quickstart} p_g^{max} \tag{18b}$$

C. Clustering Methodology

With the heterogeneity of generation units in real systems, the exact basis for clustering is a decision with important tradeoffs. Here we compare the results from four different approaches to aggregation:

1) Separate units – no clustering. This is the traditional formulation with binary commitment decisions for each unit.

2) Full clustering by unit type only – In this case all units with the same combination of fuel type and prime mover (e.g., coal steam, open cycle gas turbine, natural gas combined cycle) are combined into clusters.

3) Clustering by type and additional characteristics. This clustering approach sub-divides full clusters using an additional characteristic. For example, in this study we separately compare sub-dividing by size, age, or efficiency (heat rate). Cluster membership can be determined manually (as was done here) to provide roughly equal distributions of units between sub-clusters, or by using a formal clustering algorithm, such as *k*-means [31].

4) Clustering by plant. This approach clusters all units of the same type at the facility or plant level. Often, but not always, such units are identical.

For all clustering approaches, the representative unit for each cluster is assumed to have a size (nameplate capacity) equal to the average of cluster members. Technical characteristics such as heatrate, ramp rates, minimum output, etc., are taken as the size-weighted average. This representative plant is effectively duplicated such that the number of units in the cluster, $n_{\hat{g}}$, matches the original number of individual units.

D. Key Assumptions

In general, clustering assumes homogeneity of units within clusters. When clusters consist of identical units with constant incremental heat rates – i.e., only a single piecewise linear segment – the clustered solution exactly matches the traditional solution. For similar, but not identical, generators in the same cluster, they are assumed to have uniform technical characteristics such as minimum and maximum output levels, ramp rates, etc.

IV. Other Speedup Strategies

In addition to clustering, we explore other strategies for speeding up long-term unit commitment computations. As described in more detail below, these strategies fall into two categories: generic MILP heuristics and problem-specific simplifications. Both categories can be used with either traditional or clustered formulations and therefore offer comparisons of clustering with other strategies and methods to further speed up very large problems.

A. Mixed Integer Heuristics

As an alternative to clustering, it is also possible to use generic MILP solution tuning approaches to address the challenge of distinguishing very similar solutions that result from using binary variables to describe units with similar (or identical) operating characteristics. MILP branch-and-cut (and other combinatorial optimization), can waste considerable time finding and attempting to improve on such (nearly) equivalent solutions. Specific approaches include:

- The ϵ-optimal heuristic, informally known as "cheat," that can improve solution times during the branch-and-bound phase of branch-and-cut by only considering branches of the node tree that have the potential to improve the solution by more than ϵ. [24],
- Perturbing key parameters (variable cost) for truly identical units, to introduce small artificial differences and
- Imposing a merit order to help structure the problem by ensuring certain units always start before others, unless minimum up/down time constraints would be violated [5].

B. Problem-specific simplifications

In addition, simplifying assumptions to long-term unit commitment can reduce the solution time. These include:

- Constant incremental heat rate with offset: replace piecewise fuel-use with a single linear segment [11];
- Relaxed integer constraints for units with low min outputs: use relaxed commitment states for small units or units with small minimum output levels [25];
- Combined Reserves: aggregate reserve classes into three – off-line (tertiary), flexibility up, and flexibility down, similar to [26]; and
- Limited start-ups per time: replace minimum up/downtime with constraint on total startups per unit.

V. Experimental Setup

A. Overview

We solve the unit commitment problem for an example power system to compare the computation time and results of clustering versus a traditional, binary formulation. Further comparisons are made for both formulations in conjunction with speedup strategies described in Section IV.

B. Metrics of comparison

To provide results relevant to a range of applications, we compute multiple comparison metrics. In all cases, we compare experimental runs to the full traditional binary unit commitment formulation, indicated with subscript "baseline":

1) Total Cost is the objective function value for the optimization and includes all operations costs. For comparison, we report the percent difference computed as:

$$\Delta C^{total} = (C^{total} - C_{baseline}^{total}) / C_{baseline}^{total} \tag{19}$$

2) CO_{2e} Emissions are computed system-wide based on fuel usage for both operations and startup. A scalar percent difference is computed in the same manner as total cost.

3) Energy Mix is based on total annual production by generator class divided in the same way as for clustering. The energy mix for each class is computed by summing the product of power output and duration for all time periods and dividing by the total system energy production:

$$E_{\hat{g}}^{fraction} = \left(\sum_{t\in \boldsymbol{T}} P_{\hat{g},t} \cdot 1hr\right) \Bigg/ \left(\sum_{t\in \boldsymbol{T}} D_t \cdot 1hr\right) \qquad (20)$$

The mean absolute difference of this vector gives:

$$\Delta E^{mix} = \underset{\hat{g}\in \boldsymbol{G}}{\text{Mean}} \left| E_{\hat{g}}^{fraction} - E_{\hat{g},baseline}^{fraction} \right| \qquad (21)$$

4) Commitment Plan differences are first computed as an array of differences with one element for each time period for each group of units aggregated to the cluster level. Two scalar comparisons are then made: a) The total count of differences between plans, computed as the number of non-zero elements in this array and b) the normalized mean absolute difference where commitment difference values for each time are normalized based on the total number of units committed for that time period in the baseline:

$$\Delta U = \underset{t\in \boldsymbol{T},\hat{g}\in \boldsymbol{G}}{\text{Mean}} \left| \frac{U_{\hat{g},t} - U_{\hat{g},t}\big|_{baseline}}{\sum_{\hat{g}\in \boldsymbol{G}} \left(U_{\hat{g},t}\big|_{baseline} \right)} \right| \qquad (22)$$

5) Hourly Power Output differences are computed identically to commitment, except that for the count of differences, power levels are first rounded to the nearest 0.5MW.

6) Computation Time is reported as total solver (CPLEX) run time and excludes GAMS setup and output processing.

C. Implementation

All runs share a common model written in GAMS [35] that uses pre-compile flags for different data, model simplifications and solver configurations. The resulting problems were then solved using the state-of-the-art CPLEX 12.2 mixed-integer solver [36]. The solver was instructed to conserve memory when possible (memoryemphasis=1) to prevent out-of-memory errors for larger problems. The linear programming (LP) tolerance (epopt) was tightened to 1e-9 to ensure that the final LP solve matches the MILP branch-and-cut solution. The solver time limit (reslim) was set to 10 hours.

All model runs were conducted as a single thread running on a single 64-bit core (Intel Nehalem) at 2.67GHz clock speed. Up to 6 runs were run in parallel as sub-tasks of exclusive jobs on identical 8-core machines (2+ cores idle) with 24GB of shared RAM. Although run on a high performance cluster, the resulting resources allocated to each run are roughly equivalent to a modern personal computer.

VI. Test System #1: IEEE Reliability Test System

A. System Description

The basic IEEE Reliability Test System (RTS) was initially defined in 1979 [37], updated in 1986 [38], and again in 1996 [39]. It includes detailed unit data for ten types of generators, with between one and six units of each type, to describe a basic system with 32 total units. The dataset includes tables for demand dynamics up to a full year at an hourly resolution.

For our analysis, we use the 1996 revision [39] for demand data and most unit data including heat rates, minimum up/down times, cycling, ramping, emissions, and startup fuel usage. For each unit, our baseline formulation uses a three segment piecewise linear fuel use function with intersections at each of the provided net heat rate data points. Unit cost data is only reported in the 1979 definition [29]. System reserve requirements were taken as 1% of the load for regulation up and down, and 2% of load for load following up and down plus spinning reserves equal to the largest single unit, 400MW. Quick start reserves are not used.

We simplified the system by ignoring transmission and assuming all units are located a single node. We also removed the six hydro units leaving 26 units of eight different types. To compensate for the removed hydro, demand data was scaled uniformly by 92%, the annual ratio of hydro energy to total demand. Runs were conducted using data for the peak week.

B. Clustering Approach

The inherent duplication of identical units in the IEEE RTS system provides straightforward clustering by grouping identical units. For perturbed runs, each unit's variable operations and maintenance costs are adjusted slightly (up to 0.01%).

C. Results

Mixed Integer Heuristics As seen in Fig. 2, most of the heuristics for streamlining similar MIP results can provide some computational speedup, but only clustering provides speed up in all cases. Clustering is also significantly more effective than the other techniques, providing 100-10,000 times faster performance than the next closest heuristic. In all cases, the aggregate errors are minimal: below 0.1% for most metrics, with the only exception of approximately 0.3% errors for separate units with a 0.1% relative cheat.

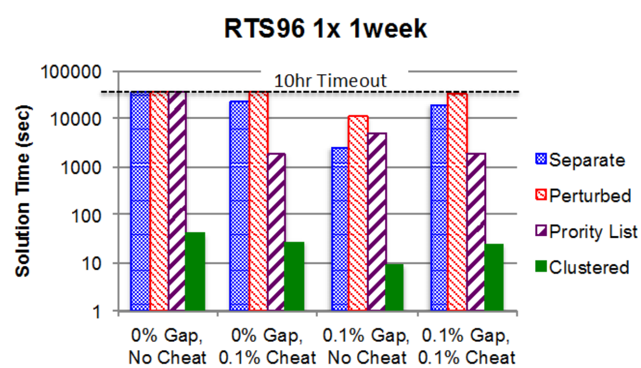


Fig. 2: MIP heuristic comparison for IEEE Reliability Test System 1996 showing solver run times for different heuristic combinations (note logarithmic time axis). In all cases, the "Separate, 0% MIP gap, No Cheat" configuration was used as a baseline.

1) Unit Commitment Simplifications: As seen in Fig. 3, most of the unit commitment simplifications also provide some speed up to the problem, with reasonably small aggregate errors, but none are as effective as clustering alone, which is about 20 times faster than any other approximation. Furthermore, the combination of clustering with the other simplifications provided additional speed-ups of 20 to over 250 times, while still maintaining small errors. With all simplifications, the aggregate errors are all less than 1.5%, and most are below 0.25% with the exception of normalized commitment errors from 0.75% to 1.5% for combined reserves and runs without minimum up and down times. And total cost errors around 0.6% for runs without minimum up/down times. Complete result tables for these and other figures, as well as additional run configurations are provided in the appendix.

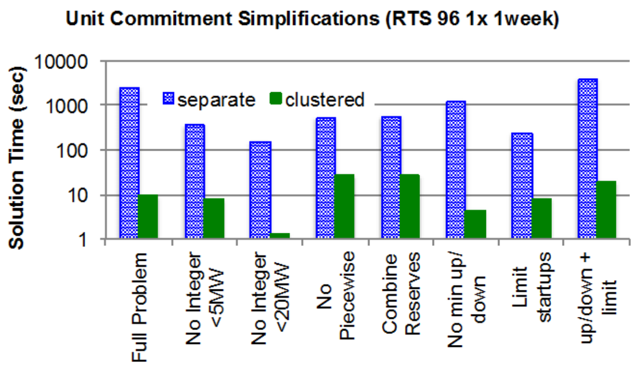


Fig. 3: Unit Commitment simplification comparison for IEEE Reliability Test System 1996 showing solver run times for different simplifications. Note logarithmic time axis. All runs used a MIP gap of 0.1% and no cheat.

VII. Test System #2: ERCOT

A. System Description

To test the impact of clustering on a system of realistic size, we modeled the entire Electric Reliability Council of Texas (ERCOT) balancing area using hourly historic demand and wind data from 2007. This system includes the entire Texas Interconnect, which covers the majority of the state of Texas and has negligible power exchange with other systems. ERCOT had a 2007 peak load of 62GW [37] supplied by a total of 92.5GW of generation capacity from 672 units [38].

To simplify the problem, we ignored the non-dispatchable combined heat and power facilities (15GW in 204 units), hydro (an additional 0.5GW in 41 units), units with uncommon fuel types (an additional 0.1GW in 72 units), and units with less than 50MW nameplate capacity (1GW in 56units). In addition, we model combined cycle facilities as 36 groups instead of 115 individual combustion and steam turbines. This resulted in a total of 205 units in our model system. We also ignore wind expansion during the year and assume a fixed wind capacity equal to the final 2007 capacity of 3.7GW. Hourly wind production was taken as this capacity times the actual percent production based on the installed capacity in each time period. Historic hourly wind production and demand data from 2007 was obtained from ERCOT [37].

The week of Saturday Mar 17, 2007 was used for 1-week (168hr) analysis. This week contains both the peak wind and minimum demand. Thirteen week data include this peak wind week plus one week for each month.

Plant-level heat rate and unit nameplate (maximum) capacity data was taken from eGrid 2010 v1.1, which contains 2007 emission and plant data. Additional generator technical parameters were taken from the Sixth Northwest Power Plan appendix I [39] for corresponding plant types. Fuel costs were based on EIA 2007 data for south central west electric power sector use [40]. Reserve requirements were taken as 1% of load for regulation up and down, 1350MW for spinning reserves, and 2% of load for load following up and down. As a simple proxy for additional reserves required for wind uncertainty, load following requirements were increased as a function of both installed capacity and wind production using the factors in [33]. Up to 50% of the spinning reserve and load following up requirements can be met by quick start open cycle natural gas units.

Complete generator data tables are provided in [34]. Hourly demand and wind profile data is available by request from ERCOT. Based on the results in [34], we used no cheat with a 0.1% MIP gap for all runs.

B. Clustering Approach

We compared the four clustering approaches described in Section III-B. The resulting number of clusters and example problem sizes are included in Table I.

TABLE I:
Size and times for 1-week (168 hr) ERCOT case.

	Problem size (Before CPLEX pre-solve)					time
aggregation	clusters	equations	variables	discretes	non-zeros	(sec)
separate	205	446,394	349,960	34,272	2,068,949	4517.2
by plant	90	197,922	151,048	14,952	943,685	435.3
by size	17	37,650	27,400	2,688	186,173	10.2
full cluster	7	14,802	10,264	1,008	74,957	2.2

C. Results

Unit Commitment simplifications. As seen in Fig. , the unit commitment simplifications provide some performance improvement. With separate units, combined reserves and constraining the number of startups, rather than using the minimum up and down time, provided the most significant speed-up of around 10 times faster calculation. But, as before, none of the simplifications were as effective as clustering alone, which was 200 times faster than any other simplification. In all cases, clustering further reduced computation time by a factor of between 350 to more than 2000. Errors for the various metrics were minimal, below 0.5% for separate units and near or below 1% for clusters.[2] The only exception was with separate units and combined reserves where normalized commitment error rose to 2.3%. The heterogeneity in generator characteristics results in relatively large (~1.25%) errors in CO_2 emissions with full clustering. The CO_2 errors are notably reduced with less aggregated clustering (next section) and longer model periods (not shown). In capacity expansion applications, future hypothetical units would be more homogenous by technology, and these errors would likely be smaller.

1) Cluster Comparison. Fig. 4 shows how most sub-clustering schemes result in small errors (around or below 1%) with the exception of clustering by age, with larger errors (2.3-4.5%) for all metrics except CO_2 emissions. Clustering by

[2] Unlike operational unit commitment where a 1% cost savings represents a major advance, here a few percent error is in line with other expected errors.

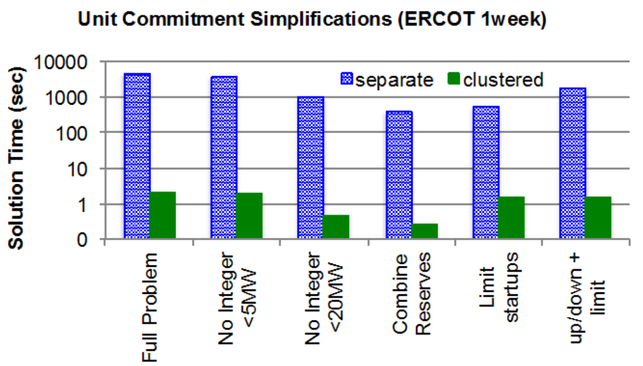


Fig. 4: Unit Commitment comparison of solver run times for different simplifications for ERCOT 2007. Note logarithmic time axis. The full problem with separate units was used as a baseline.

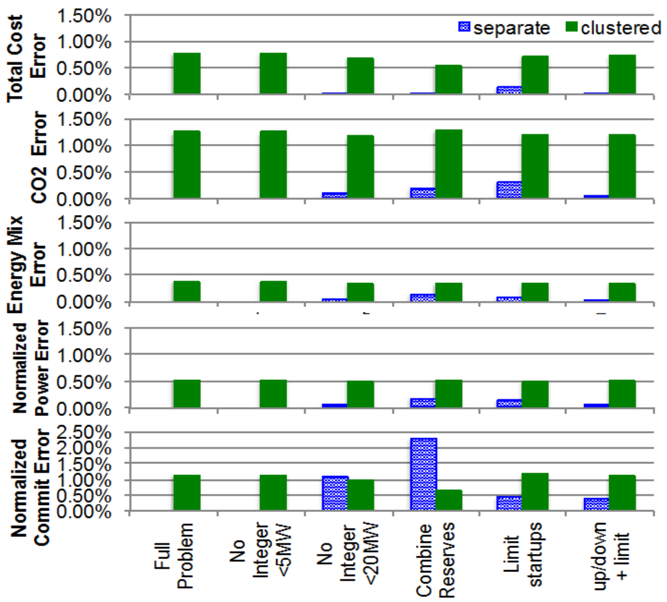


Fig. 3: Unit Commitment comparison of key error metrics for ERCOT 2007 The full problem with separate units was used as a baseline.

efficiency resulted in the lowest errors among the 17-cluster runs, for all other metrics, often close to or slightly better than the larger clustering by plant formulation (90 units).

2) Cluster Scaling. Fig. 5 shows how the total solver time is greatly reduced by clustering, enabling tractable computation of a full year, 8760 hour, optimal unit commitment for both 17 clusters (less than 3 hours) and 7 clusters (130 seconds). The primary driver for these speed-ups is a drastic reduction in the numbers of variables and equations which both scale roughly proportionally to the number of clusters.

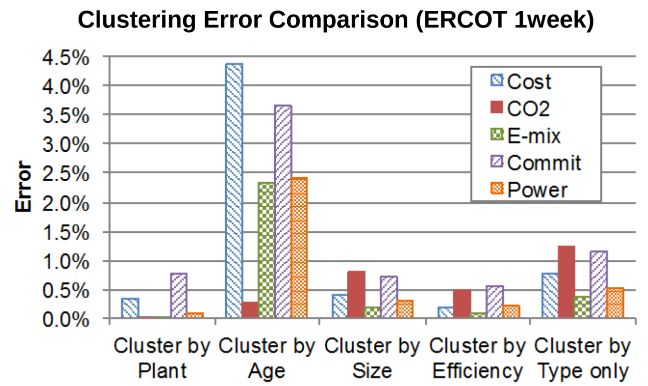


Fig. 4: Error comparison for different clustering approaches for ERCOT 2007. In all cases, the full problem with separate units was used as a baseline. All runs used a MIP gap of 0.1% and no cheat.

VIII. Conclusions

In this paper, we demonstrated the tradeoff between accuracy and runtime resulting from different levels of aggregation for heterogeneous clusters and other heuristic simplifications in unit commitment. In comparison to traditional binary formulations, clustering provides orders of magnitude faster computation – from 10 to over 1000 times faster depending on the configuration – by grouping similar units into clusters and assigning an integer, rather than binary, commitment decision to the group. This assumption builds on the existing concept of aggregating identical units. Clustering allows capturing full unit commitment constraints – including ramping, startup costs, minimum output levels, and minimum up and down times – at an individual unit level under the key assumption that all units with in a cluster are identical. Despite this assumption, we show that errors are small for a wide range of metrics.

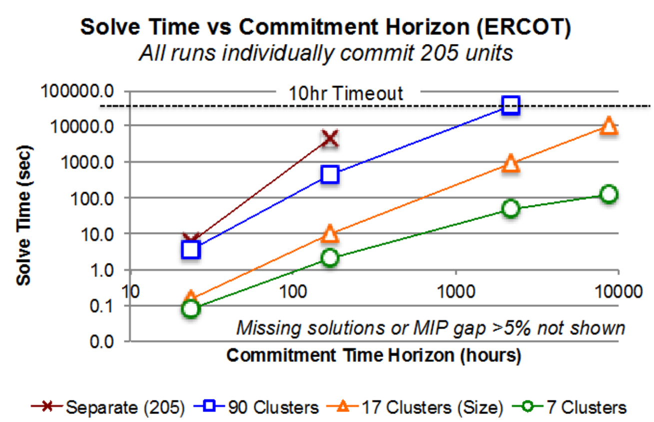


Fig. 5: Impact of clustering and model time horizon on solution time. Note both axes are logarithmic. All runs conducted with a 0.1% MIP gap and no cheat. Due to data limitations, constant heat rates are assumed. No other simplifications were used.

A numeric example using an ERCOT-based 205-unit system shows that careful aggregation (17 clusters) introduces errors of 0.05-0.2% for total cost, CO_2 emissions, energy mix, and dispatch schedule while providing several orders of magnitude faster solution times (400x) compared to traditional binary formulations. The unit commitment metric exhibits errors of around 0.9%. More aggressive aggregation (seven clusters) increases errors somewhat roughly double) but achieves further speedup (2000x). We also demonstrate a full year (8760 hour) unit commitment for a 205-unit system in less than three minutes with personal computer hardware.

We also compare other unit commitment simplifications – notably combining reserves and relaxing integer constraints for units with small minimum output levels – that can provide an additional order of magnitude speed-up for some problems. The clustering approach demonstrated here provides the ability to capture unit-level commitment decisions with intertemporal (hourly) constraints within a single optimization problem, which can be embedded within longer term operational and strategic analyses such as hydro-thermal coordination or capacity expansion under emissions or other policy constraints, especially when the long-term problem is stochastic. The application of clustering to capacity expansion and other long-term strategic decisions are left for future research.

IX. Acknowledgements

The authors wish to thank A.J. Conejo, J.I. Perez-Arriaga, and C. Batlle for valuable feedback on this manuscript.

X. References

[1] T. S. Dillon, K. W. Edwin, H.-D. Kochs, and R. J. Taud, "Integer Programming Approach to the Problem of Optimal Unit Commitment with Probabilistic Reserve Determination," *IEEE Trans. Power App. Syst*., vol. PAS-97, no. 6, pp. 2154-2166, Nov. 1978.

[2] J.M. Arroyo and A.J. Conejo, "Optimal response of a thermal unit to an electricity spot market," *IEEE Trans. Power Syst*., vol. 15, no. 3, pp. 1098-1104, Aug 2000.

[3] J. Ostrowski, M. F. Anjos, and A. Vannelli, "Tight Mixed Integer Linear Programming Formulations for the Unit Commitment Problem," *IEEE Trans. Power Syst*., vol. 27, no. 1, pp. 39-46, Feb. 2012.

[4] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit Commitment for Systems With Significant Wind Penetration," *IEEE Trans. Power Syst*., vol. 24, no. 2, pp. 592-601, May 2009.

[5] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve Requirements for Wind Power Integration: A Scenario-Based Stochastic Programming Framework," *IEEE Trans. Power Syst*., vol. 26, no. 4, pp. 2197-2206, Nov. 2011.

[6] J. M. Morales, A. J. Conejo, and J. Perez-Ruiz, "Economic Valuation of Reserves in Power Systems With High Penetration of Wind Power," *IEEE Trans. Power Syst*., vol. 24, no. 2, pp. 900-910, May 2009.

[7] F. Bouffard and F. D. Galiana, "Stochastic Security for Operations Planning With Significant Wind Power Generation," *IEEE Trans. Power Syst*., vol. 23, no. 2, pp. 306–316, May 2008.

[8] V. S. Pappala, I. Erlich, K. Rohrig, and J. Dobschinki, "A Stochastic Model for the Optimal Operation of a Wind-Thermal Power System," *IEEE Trans. Power Syst*., vol. 24, no. 2, pp. 940–950, May 2009.

[9] E. M. Constantinescu, V. M. Zavala, M. Rocklin, S. Lee, and M. Anitescu, "A Computational Framework for Uncertainty Quantification and Stochastic Optimization in Unit Commitment With Wind Power Generation," *IEEE Trans. Power Syst*., vol. 26, no. 1, pp. 431-441, Feb. 2011.

[10] P. Meibom, R. Barth, B. Hasche, H. Brand, C. Weber, and M. O'Malley, "Stochastic Optimization Model to Study the Operational Impacts of High Wind Penetrations in Ireland," *IEEE Trans. Power Syst*., vol. 26, no. 3, pp. 1367-1379, Aug. 2011.

[11] J. J. Hargreaves and B. F. Hobbs, "Commitment and Dispatch With Uncertain Wind Generation by Dynamic Programming," *IEEE Trans. Sust. Energy*, vol. 3, no. 4, pp. 724-734, Oct. 2012.

[12] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem," *IEEE Trans. Power Syst*., (in press).

[13] M. A. Ortega-Vazquez and D. S. Kirschen, "Estimating the Spinning Reserve Requirements in Systems With Significant Wind Power Generation Penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 114–124, Feb. 2009.

[14] J. F. Restrepo and F. D. Galiana, "Assessing the Yearly Impact of Wind Power Through a New Hybrid Deterministic/Stochastic Unit Commitment," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 401–410, Feb. 2011.

[15] B. Palmintier and M. Webster, "Impact of Unit Commitment Constraints on Generation Expansion Planning with Renewables," in *Proceedings of 2011 IEEE Power and Energy Society General Meeting*, Detroit, MI, 2011.

[16] S. Kamalinia and M. Shahidehpour, "Generation expansion planning in wind-thermal power systems," *IET Generation, Transmission & Distribution*, vol. 4, no. 8, p. 940, 2010.

[17] D. S. Kirschen, J. Ma, V. Silva, and R. Belhomme,, "Evaluating and Planning Flexibility in Sustainable Power Systems," *IEEE Trans. Sust. Energy*, vol. 4, no. 1, pp. 200-209, Jan. 2013.

[18] N. Langrene, W. van Ackooij, and F. Breant, "Dynamic Constraints for Aggregated Units: Formulation and Application," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1349–1356, Aug. 2011.

[19] S. Cerisola, Á. Baíllo, J. M. Fernández-López, A. Ramos, and R. Gollmer, "Stochastic Power Generation Unit Commitment in Electricity Markets: A Novel Formulation and a Comparison of Solution Methods," *Oper. Res.*, vol. 57, pp. 32–46, Jan. 2009.

[20] R. Gollmer, M. P. Nowak, W. Römisch, and R. Schultz, "Unit commitment in power generation–a basic model and some extensions," *Annals of Operations Research*, vol. 96, no. 1, pp. 167–189, 2000.

[21] K. Hara, M. Kimura, and N. Honda, "A Method for Planning Economic Unit Commitment and Maintenance of Thermal Power Systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-85, no. 5, pp. 427–436, May 1966.

[22] J. Garcia-Gonzalez, R. M. R. de la Muela, L. M. Santos, and A. M. Gonzalez, "Stochastic Joint Optimization of Wind Generation and Pumped-Storage Units in an Electricity Market," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 460–468, May 2008.

[23] H. D. Sherali and J. C. Smith, "Improving Discrete Model Representations via Symmetry Considerations," *Management Science*, vol. 47, no. 10, pp. 1396 –1407, Oct. 2001.

[24] S. Cerisola, "Benders Decomposition for Mixed Integer Problems: Application to a Medium Term Hydrothermal Coordination Model," PhD Dissertation, Instituto de Investigación Tecnológica, Universidad Pontificia Comillas, Madrid, Spain, 2004.

[25] S. Sen and D. P. Kothari, "An Equivalencing Technique For Solving The Large-Scale Thermal Unit Commitment Problem," in *The Next Generation of Electric Power Unit Commitment Models*, B. F. Hobbs, M. H. Rothkopf, R. P. O'Neill, and H. Chao, Eds. Boston: Kluwer Academic Publishers, 2002, pp. 211–225.

[26] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. Wiley-Interscience, 1996.

[27] R. Baldick, "The generalized unit commitment problem," *IEEE Transactions on Power Systems*, vol. 10, no. 1, pp. 465–475, Feb. 1995.

[28] L. Wu, M. Shahidehpour, and T. Li, "Stochastic Security-Constrained Unit Commitment," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 800–811, May 2007.

[29] D. Rajan and S. Takriti, "Minimum Up/Down Polytopes of the Unit Commitment Problem with Start-Up Costs," IBM T.J. Watson Research Center, Yorktown Heights, NY, IBM Research Report, Jun. 2005.

[30] K. W. Hedman, R. P. O'Neill, and S. S. Oren, "Analyzing valid inequalities of the generation unit commitment problem," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1–6.

[31] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, Jun. 2010.

[32] P. A. Ruiz, C. R. Philbrick, and P. W. Sauer, "Modeling Approaches for Computational Cost Reduction in Stochastic Unit Commitment Formulations," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 588–589, Feb. 2010.

[33] C. De Jonghe, E. Delarue, R. Belmans, and W. D'haeseleer, "Determining optimal electricity technology mix with high level of wind power penetration," *Applied Energy*, vol. 88, no. 6, pp. 2231–2238, Jun. 2011.

[34] B. Palmintier and M. Webster, "Heterogeneous Unit Clustering for Efficient Operational Flexibility Modeling - Extended Working Paper," 2012. Available: http://web.mit.edu/b_p/pdf/Palmintier&Webster-2012-ClusteringForFlexibility-WP.pdf

[35] GAMS Development Corporation, *General Algebraic Modeling System (GAMS)*. 2010.

[36] IBM ILOG, *CPLEX Optimizer*. 2010.

[37] Reliability Test System Task Force, “IEEE Reliability Test System,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, pp. 2047–2054, Nov. 1979.
[38] R. N. Allan, R. Billinton, and N. M. K. Abdel-Gawad, “The IEEE Reliability Test System - Extensions to and Evaluation of the Generating System,” *IEEE Transactions on Power Systems*, vol. 1, no. 4, pp. 1–7, Nov. 1986.
[39] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh, “The IEEE Reliability Test System-1996. A report prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee,” *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.
[40] Electricity Reliability Council of Texas, http://www.ercot.com/.
[41] US Environmental Protection Agency (EPA), *Emissions & Generation Resource Integrated Database (eGrid)*. 2010.
[42] Northwest Power and Conservation Council, “Sixth Northwest Conservation and Electric Power Plan,” 2010-09, Feb. 2010.
[43] US Energy Information Administration [EIA], “Annual Energy Outlook 2010, Table 3,” 2010. [Online]. Available: http://www.eia.gov/oiaf/aeo/tablebrowser/#release=AEO2010&table=3-AEO2010. [Accessed: 05-Nov-2011].

XI. Biographies

Bryan Palmintier (M ’07) has a degree of Engineer in mechanical engineering (2004) and an M.S. in aero/astro engineering (2000) from Stanford University and a B.S. in aerospace engineering (1996) from Georgia Institute of Technology. He is currently a PhD candidate in the Engineering Systems Division (ESD) at the Massachusetts Institute of Technology (MIT). His research on the electric power system aims to tractably integrate operations constraints and dynamics into long-term stochastic planning models.

Mort Webster (M ’10) holds a Ph.D. from the Engineering Systems Division (2000) at the Massachusetts Institute of Technology. Professor Webster holds an M.S. from MIT’s Technology and Policy Program (1996), and a B.S.E. in Computer Science and Engineering (1988) from the University of Pennsylvania. He is an Assistant Professor of Engineering Systems at MIT, Cambridge, MA, USA. His research and teaching focuses on environmental and energy systems analysis and decision-making under uncertainty.

I. APPENDIX

A. IEEE RTS Piecewise Linear Fuel Use

Based on data provided by IEEE RTS 1996 [31]

*	seg1.slope [MMBTU/MWh]	seg1.intercept [MMBTU]	seg2.slope [MMBTU/MWh]	seg2.intercept [MMBTU]	seg3.slope [MMBTU/MWh]	seg3.intercept [MMBTU]
Oil_St_12MW	10.155	14.068	10.900	9.600	12.400	-4.800
Oil_CT_20MW	10.023	79.632	12.395	41.684	14.400	1.980
Coal_St_76MW	9.657	113.240	10.672	74.683	12.400	-30.400
Oil_St_100MW	8.401	114.950	9.065	81.733	9.652	34.800
Coal_St_155MW	8.386	155.068	8.713	124.620	9.128	73.160
Oil_St_197MW	8.590	148.932	9.026	97.397	9.424	34.672
Coal_St_350MW	8.640	218.400	9.067	121.333	9.500	0
Nuke_400MW	8.899	385.200	9.078	349.333	9.320	272.000

B. ERCOT 2007 (Simplified) Unit Data

Adapted from eGrid 2010 v1.1 [33]

	type	**Capacity**	**fuel**	**heat rate**
Unit Name	[code]	[MW]	[name]	[MBTU/MWh]
Arthur_Von_Rosenberg_Combined	ng_cc	550	ng	7.499
Barney_M_Davis_1	ng_st	352	ng	11.415
Barney_M_Davis_2	ng_st	351	ng	11.415
Bastrop_Combined	ng_cc	727.8	ng	7.845
Big_Brown_1	coal_lig_st	593.4	coal_lig	10.698
Big_Brown_2	coal_lig_st	593.4	coal_lig	10.698
Bosque_County_Peaking_GT_1	ng_gt	154	ng	7.639
Bosque_County_Peaking_GT_2	ng_gt	154	ng	7.639
Bosque_County_Peaking_Units3to5_Combined	ng_cc	499	ng	7.639
Brazoz_Valley_Generating_Facility_Combined	ng_cc	675.6	ng	7.462
Bryan_6	ng_st	54	ng	21.683
Cedar_Bayou_1	ng_st	765	ng	10.729
Cedar_Bayou_2	ng_st	765	ng	10.729
Coleto_Creek_1	coal_sub_st	600.4	coal_sub	10.133
Colorado_Bend_Energy_Center_Combined_1	ng_cc	278.1	ng	7.386
Colorado_Bend_Energy_Center_Combined_2	ng_cc	278.1	ng	7.386
Comanche_Peak_1	u235_st	1215	u235	10.400
Comanche_Peak_2	u235_st	1215	u235	10.400
Dansby_1	ng_st	105	ng	11.288
Decker_Creek_1	ng_st	321	ng	11.002
Decker_Creek_2	ng_st	405	ng	11.002
Decker_Creek_GT1	ng_gt	51.5	ng	11.002
Decker_Creek_GT2	ng_gt	51.5	ng	11.002
Decker_Creek_GT3	ng_gt	51.5	ng	11.002
Decker_Creek_GT4	ng_gt	51.5	ng	11.002
DeCordova_Steam_Electric_Station_1	ng_st	799.2	ng	12.147
DeCordova_Steam_Electric_Station_CT1	ng_gt	89.4	ng	12.147
DeCordova_Steam_Electric_Station_CT2	ng_gt	89.4	ng	12.147
DeCordova_Steam_Electric_Station_CT3	ng_gt	89.4	ng	12.147
DeCordova_Steam_Electric_Station_CT4	ng_gt	89.4	ng	12.147
Ennis_Power_Company_Combined	ng_cc	418	ng	7.361
Exelon_LaPorte_Generating_Station_GT1	ng_gt	59	ng	12.676
Exelon_LaPorte_Generating_Station_GT2	ng_gt	59	ng	12.676
Exelon_LaPorte_Generating_Station_GT3	ng_gt	59	ng	12.676
Exelon_LaPorte_Generating_Station_GT4	ng_gt	59	ng	12.676
Fayette_Power_Project_1	coal_sub_st	615	coal_sub	10.679
Fayette_Power_Project_2	coal_sub_st	615	coal_sub	10.679
Fayette_Power_Project_3	coal_sub_st	460	coal_sub	10.679
Forney_Energy_Center_Combined_1	ng_cc	891.9	ng	7.351
Forney_Energy_Center_Combined_2	ng_cc	891.9	ng	7.351
Freestone_Power_Generation_LP_Combined_1	ng_cc	518	ng	7.522
Freestone_Power_Generation_LP_Combined_2	ng_cc	518	ng	7.522
Frontera_Energy_Center_Combined	ng_cc	529	ng	7.535

[1]his work was supported in part by the U.S. National Science Foundation under Grants 1128147 and 0835414.

B. S. Palmintier (e-mail: b_p@mit.edu) and M. D. Webster (e-mail: mort@mit.edu) are with the Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Gibbons_Creek_1	coal_sub_st	453.5	coal_sub	9.977
Graham_1	ng_st	247.7	ng	11.947
Graham_2	ng_st	387	ng	11.947
Greens_Bayou_5	ng_st	446.4	ng	14.681
Greens_Bayou_73	ng_gt	72	ng	14.681
Greens_Bayou_74	ng_gt	72	ng	14.681
Greens_Bayou_81	ng_gt	72	ng	14.681
Greens_Bayou_82	ng_gt	72	ng	14.681
Greens_Bayou_83	ng_gt	72	ng	14.681
Greens_Bayou_84	ng_gt	72	ng	14.681
Guadalupe_Generating_Station_Combined_1	ng_cc	571.1	ng	7.423
Guadalupe_Generating_Station_Combined_2	ng_cc	571.1	ng	7.423
Handley_2	ng_st	74.8	ng	13.823
Handley_3	ng_st	404.8	ng	13.823
Handley_4	ng_st	455	ng	13.823
Handley_5	ng_st	455	ng	13.823
Hays_Energy_Project_U1	ng_cc	241.7	ng	7.158
Hays_Energy_Project_U2	ng_cc	241.7	ng	7.158
Hays_Energy_Project_U3	ng_cc	252.8	ng	7.158
Hays_Energy_Project_U4	ng_cc	252.8	ng	7.158
Hidalgo_Energy_Center_Combined	ng_cc	551.3	ng	7.219
J_K_Spruce_1	coal_sub_st	566	coal_sub	10.822
J_K_Spruce_2	coal_sub_st	820	coal_sub	10.822
J_T_Deely_1	coal_sub_st	486	coal_sub	14.073
J_T_Deely_2	coal_sub_st	446	coal_sub	14.073
Jack_County_Combined	ng_cc	640	ng	7.284
Kiamichi_Energy_Facility_Combined_1	ng_cc	685	ng	7.397
Kiamichi_Energy_Facility_Combined_2	ng_cc	685	ng	7.397
Lake_Creek_ST1	ng_st	79.6	ng	14.369
Lake_Creek_ST2	ng_st	236	ng	14.369
Lake_Hubbard_1	ng_st	396.5	ng	12.159
Lake_Hubbard_2	ng_st	531	ng	12.159
Lamar_Power_Project_Combined_1	ng_cc	545.4	ng	7.768
Lamar_Power_Project_Combined_2	ng_cc	545.4	ng	7.768
Laredo_3	ng_st	115.2	ng	11.592
Leon_Creek_3	ng_st	75	ng	11.834
Leon_Creek_4	ng_st	113.7	ng	11.834
Leon_Creek_CGT1	ng_gt	57.4	ng	11.834
Leon_Creek_CGT2	ng_gt	57.4	ng	11.834
Leon_Creek_CGT3	ng_gt	57.4	ng	11.834
Leon_Creek_CGT4	ng_gt	57.4	ng	11.834
Limestone_1	coal_lig_st	893	coal_lig	9.612
Limestone_2	coal_lig_st	956.8	coal_lig	9.612
Lost_Pines_1_Power_Project_Combined	ng_cc	595	ng	7.217
Magic_Valley_Generating_Station_Combined	ng_cc	801	ng	7.275
Martin_Lake_1	coal_lig_st	793.2	coal_lig	11.090
Martin_Lake_2	coal_lig_st	793.2	coal_lig	11.090
Martin_Lake_3	coal_lig_st	793.2	coal_lig	11.090
Midlothian_Energy_Facility_STK1	ng_cc	289	ng	7.460
Midlothian_Energy_Facility_STK2	ng_cc	289	ng	7.460
Midlothian_Energy_Facility_STK3	ng_cc	289	ng	7.460
Midlothian_Energy_Facility_STK4	ng_cc	289	ng	7.460
Midlothian_Energy_Facility_STK5	ng_cc	289	ng	7.460
Midlothian_Energy_Facility_STK6	ng_cc	289	ng	7.460
Monticello_1	coal_lig_st	593.4	coal_lig	10.916
Monticello_2	coal_lig_st	593.4	coal_lig	10.916
Monticello_3	coal_lig_st	793.2	coal_lig	10.916
Morgan_Creek_5	ng_st	170.4	ng	13.844
Morgan_Creek_6	ng_st	517.5	ng	13.844
Morgan_Creek_CT1	ng_gt	89.4	ng	13.844
Morgan_Creek_CT2	ng_gt	89.4	ng	13.844
Morgan_Creek_CT3	ng_gt	89.4	ng	13.844
Morgan_Creek_CT4	ng_gt	89.4	ng	13.844
Morgan_Creek_CT5	ng_gt	89.4	ng	13.844
Morgan_Creek_CT6	ng_gt	89.4	ng	13.844
Mountain_Creek_3	ng_st	74.9	ng	12.481
Mountain_Creek_6	ng_st	135.7	ng	12.481
Mountain_Creek_7	ng_st	136	ng	12.481
Mountain_Creek_8	ng_st	580.5	ng	12.481
Newgulf_Cogen_GEN1	ng_gt	78.7	ng	13.784
North_Lake_1	ng_st	176.8	ng	11.651

North_Lake_2	ng_st	170.4	ng	11.651
North_Lake_3	ng_st	361.3	ng	11.651
O_W_Sommers_1	ng_st	446	ng	12.109
O_W_Sommers_2	ng_st	446	ng	12.109
Odessa_Ector_Generating_Station_Combined_1	ng_cc	567.6	ng	7.604
Odessa_Ector_Generating_Station_Combined_2	ng_cc	567.6	ng	7.604
Oklaunion_1	coal_sub_st	720	coal_sub	10.582
P_H_Robinson_1	ng_st	484.5	ng	13.008
P_H_Robinson_2	ng_st	484.5	ng	13.008
P_H_Robinson_3	ng_st	580.5	ng	13.008
P_H_Robinson_4	ng_st	765	ng	13.008
Permian_Basin_5	ng_st	114.9	ng	13.750
Permian_Basin_6	ng_st	535.5	ng	13.750
Permian_Basin_CT1	ng_gt	89.4	ng	13.750
Permian_Basin_CT2	ng_gt	89.4	ng	13.750
Permian_Basin_CT3	ng_gt	89.4	ng	13.750
Permian_Basin_CT4	ng_gt	89.4	ng	13.750
Permian_Basin_CT5	ng_gt	89.4	ng	13.750
Quail_Run_Energy_Center_Combined_1	ng_cc	298	ng	8.540
Quail_Run_Energy_Center_Combined_2	ng_cc	275	ng	8.540
R_W_Miller_1	ng_st	66	ng	12.666
R_W_Miller_2	ng_st	100	ng	12.666
R_W_Miller_3	ng_st	200	ng	12.666
R_W_Miller_4	ng_gt	118.8	ng	12.666
R_W_Miller_5	ng_gt	118.8	ng	12.666
Ray_Olinger_1	ng_st	75	ng	12.326
Ray_Olinger_2	ng_st	113.4	ng	12.326
Ray_Olinger_3	ng_st	156.6	ng	12.326
Ray_Olinger_4	ng_gt	82.7	ng	12.326
Rio_Nogales_Power_Project_Combined	ng_cc	898.2	ng	7.298
Sam_Bertron_3	ng_st	225.3	ng	11.572
Sam_Bertron_4	ng_st	225.3	ng	11.572
Sam_Bertron_ST1	ng_st	187.8	ng	11.572
Sam_Bertron_ST2	ng_st	187.8	ng	11.572
Sam_Rayburn_Units7to10_Combined	ng_cc	189.6	ng	9.147
San_Jacinto_Steam_Electric_Station_1	ng_gt	88.2	ng	13.516
San_Jacinto_Steam_Electric_Station_2	ng_gt	88.2	ng	13.516
San_Miguel_1	coal_lig_st	410	coal_lig	12.148
Sand_Hill_5Combined	ng_cc	388	ng	7.328
Sand_Hill_SH1	ng_gt	51.4	ng	7.328
Sand_Hill_SH2	ng_gt	51.4	ng	7.328
Sand_Hill_SH3	ng_gt	51.4	ng	7.328
Sand_Hill_SH4	ng_gt	51.4	ng	7.328
Sandow_No_4_4	coal_lig_st	590.6	coal_lig	11.163
Silas_Ray_10	ng_gt	61	ng	11.083
Silas_Ray_Units6and9_Combined	ng_cc	86	ng	11.083
Sim_Gideon_1	ng_st	144	ng	11.456
Sim_Gideon_2	ng_st	144	ng	11.456
Sim_Gideon_3	ng_st	351	ng	11.456
South_Texas_Project_1	u235_st	1354.3	u235	10.400
South_Texas_Project_2	u235_st	1354.3	u235	10.400
Spencer_4	ng_st	61.1	ng	16.622
Spencer_5	ng_st	65.4	ng	16.622
Stryker_Creek_ST1	ng_st	176.8	ng	11.375
Stryker_Creek_ST2	ng_st	526.6	ng	11.375
T_H_Wharton_3Combined	ng_cc	318.3	ng	9.529
T_H_Wharton_4Combined	ng_cc	329.1	ng	9.529
T_H_Wharton_51	ng_gt	85	ng	9.529
T_H_Wharton_52	ng_gt	85	ng	9.529
T_H_Wharton_53	ng_gt	85	ng	9.529
T_H_Wharton_54	ng_gt	85	ng	9.529
T_H_Wharton_55	ng_gt	85	ng	9.529
T_H_Wharton_56	ng_gt	85	ng	9.529
Tenaska_Frontier_Generation_Station_Combined	ng_cc	939.7	ng	6.901
Tenaska_Gateway_Generation_Station_Combined	ng_cc	939.6	ng	7.475
Thomas_C_Ferguson_1	ng_st	446	ng	10.994
Tradinghouse_1	ng_st	580.5	ng	11.838
Tradinghouse_2	ng_st	799.2	ng	11.838
Trinidad_6	ng_st	239.3	ng	13.508
Twin_Oaks_Power_One_1	coal_lig_st	174.6	coal_lig	10.860
Twin_Oaks_Power_One_2	coal_lig_st	174.6	coal_lig	10.860

V_H_Braunig_1	ng_st	225	ng	11.161
V_H_Braunig_2	ng_st	252	ng	11.161
V_H_Braunig_3	ng_st	417	ng	11.161
Valley_1	ng_st	198.9	ng	13.664
Valley_2	ng_st	580.5	ng	13.664
Valley_3	ng_st	396	ng	13.664
W_A_Parish_1	ng_st	187.8	ng	10.382
W_A_Parish_2	ng_st	187.8	ng	10.382
W_A_Parish_3	ng_st	299.2	ng	10.382
W_A_Parish_4	ng_st	580.5	ng	10.382
W_A_Parish_5	coal_sub_st	734.1	coal_sub	10.382
W_A_Parish_6	coal_sub_st	734.1	coal_sub	10.382
W_A_Parish_7	coal_sub_st	614.6	coal_sub	10.382
W_A_Parish_8	coal_sub_st	614.6	coal_sub	10.382
W_B_Tuttle_1	ng_st	75	ng	17.474
W_B_Tuttle_3	ng_st	113.6	ng	17.474
W_B_Tuttle_4	ng_st	191.7	ng	17.474
Wind	wind	3710.5	wind	1.000
Wise_County_Power_LP_Combined	ng_cc	746	ng	7.609
Wolf_Hollow_I_LP_Combined	ng_cc	809.6	ng	7.882

C. ERCOT 2007 Clustering Parameters

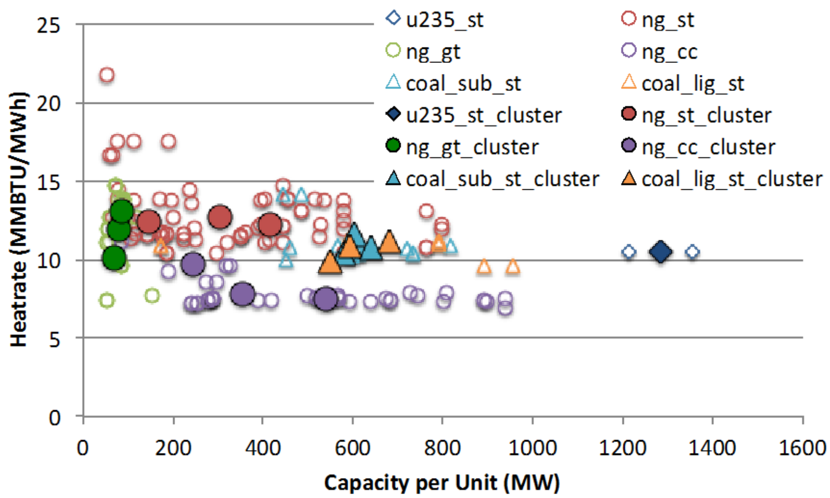
			Age (year in service)			Size (Cap >= X MW)			Efficiency (HR >= X MMBTU/MWh)		
Type	Fuel	Use Clusters	old	midAge	new	small	med	large	hiEff	avgEff	loEff
coal_lig_st	coal_lig	TRUE	0	1975	1985	0	500	650	0	10.5	12
coal_sub_st	coal_sub	TRUE	0	1980	1985	0	500	650	0	10.2	12
ng_cc	ng	TRUE	0	2000	2005	0	300	600	0	7.5	9
ng_gt	ng	TRUE	0	1980	2000	0	75	100	0	10	13
ng_st	ng	TRUE	0	1960	1970	0	150	300	0	11.5	13
u235_st	u235	FALSE									
wind	wind	FALSE									

D. ERCOT Clusters

1) Cluster by Type only

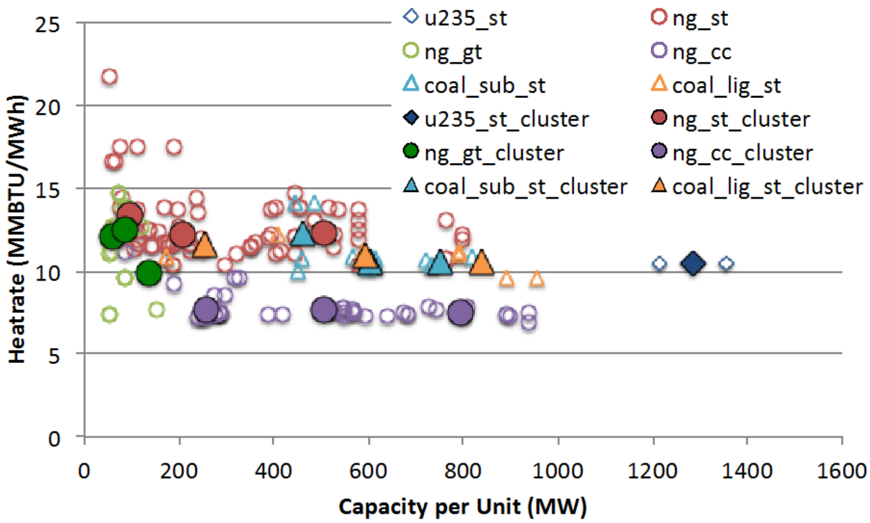
cluster_name	# units	weighted average heat rate	Average Capacity (MW)	Total Capacity (MW)
coal_lig_st	14	10.730	624.71	8745.94
coal_sub_st	14	10.897	605.66	8479.24
ng_cc	46	7.551	500.89	23040.94
ng_gt	52	11.969	78.93	4104.36
ng_st	74	12.340	309.34	22891.16
u235_st	4	10.400	1284.65	5138.60
wind	1	1.000	3710.50	3710.50

2) Cluster by Age



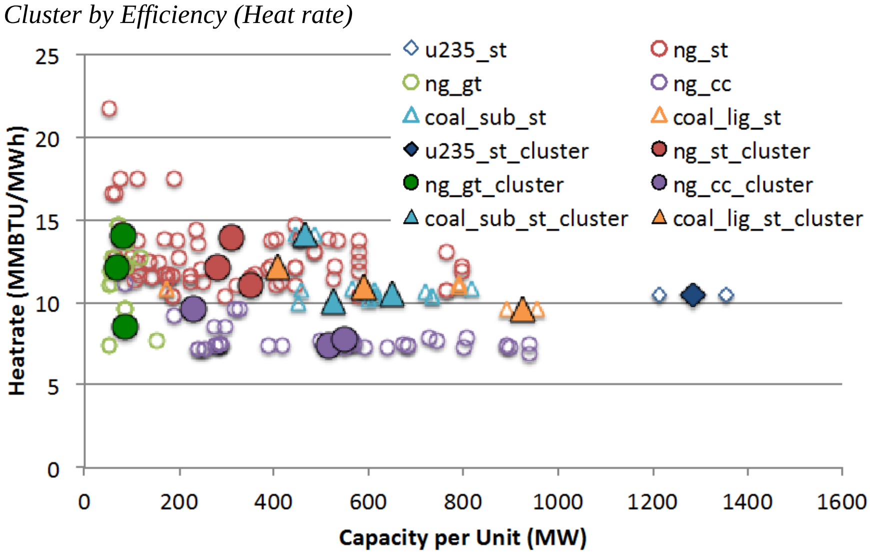
cluster_name	# units	weighted average heat rate	Average Capacity (MW)	Total Capacity (MW)
coal_lig_st_old	3	10.770	593.40	1780.20
coal_lig_st_midAge	7	11.139	680.97	4766.79
coal_lig_st_new	4	9.810	549.75	2199.00
coal_sub_st_old	5	11.584	603.04	3015.20
coal_sub_st_midAge	5	10.330	579.62	2898.10
coal_sub_st_new	4	10.729	641.50	2566.00
ng_cc_old	3	9.711	244.47	733.41
ng_cc_midAge	38	7.459	540.48	20538.24
ng_cc_new	5	7.723	353.84	1769.20
ng_gt_old	12	11.891	78.50	942.00
ng_gt_midAge	24	13.071	84.99	2039.76
ng_gt_new	16	10.031	70.18	1122.88
ng_st_old	18	12.347	145.34	2616.12
ng_st_midAge	27	12.624	303.53	8195.31
ng_st_new	29	12.146	416.55	12079.95
u235_st	4	10.400	1284.65	5138.60
wind	1	1.000	3710.50	3710.50

3) Cluster by Size



cluster_name	# units	weighted average heat rate	Average Capacity (MW)	Total Capacity (MW)
coal_lig_st_small	3	11.556	253.07	759.21
coal_lig_st_med	5	10.878	592.84	2964.20
coal_lig_st_large	6	10.518	837.10	5022.60
coal_sub_st_small	4	12.221	461.38	1845.52
coal_sub_st_med	6	10.510	604.27	3625.62
coal_sub_st_large	4	10.550	752.05	3008.20
ng_cc_small	16	7.681	257.99	4127.84
ng_cc_med	17	7.651	504.82	8581.94
ng_cc_large	13	7.415	794.72	10331.36
ng_gt_small	23	12.042	59.57	1370.11
ng_gt_med	25	12.456	87.55	2188.75
ng_gt_large	4	9.828	136.40	545.60
ng_st_small	21	13.323	96.97	2036.37
ng_st_med	20	12.171	207.13	4142.60
ng_st_large	33	12.262	506.44	16712.52
u235_st	4	10.400	1284.65	5138.60
wind	1	1.000	3710.50	3710.50

4) Cluster by Efficiency (Heat rate)



cluster_name	# units	weighted average heat rate	Average Capacity (MW)	Total Capacity (MW)
coal_lig_st_hiEff	2	9.612	924.90	1849.80
coal_lig_st_avgEff	11	10.959	589.65	6486.15
coal_lig_st_loEff	1	12.148	410.00	410.00
coal_sub_st_hiEff	2	10.066	526.95	1053.90
coal_sub_st_avgEff	10	10.575	649.34	6493.40
coal_sub_st_loEff	2	14.073	466.00	932.00
ng_cc_hiEff	29	7.331	516.26	14971.54
ng_cc_avgEff	13	7.746	549.72	7146.36
ng_cc_loEff	4	9.595	230.75	923.00
ng_gt_hiEff	12	8.518	85.30	1023.60
ng_gt_avgEff	20	12.069	70.53	1410.60
ng_gt_loEff	20	13.998	83.53	1670.60
ng_st_hiEff	20	10.976	350.09	7001.80
ng_st_avgEff	29	12.028	279.65	8109.85
ng_st_loEff	25	13.892	311.20	7780.00
u235_st	4	10.400	1284.65	5138.60
wind	1	1.000	3710.50	3710.50

E. Assumed ERCOT Unit Technical Constraints

Based on Northwest Power Plan v6 [34]

type	c_var_om	c_fix_om	ramp_max	unit_min_pu	fuel_start	c_start_fix	max_start	quick_start	reg_up	reg_down	spin_rsv	min_up	min_down
[code]	[$/MWh]	[$/MW-yr]	[p.u./hr]	[fract of max]	[MBTU/start]	[$/start]	[start/yr]	[p.u.]	[p.u.]	[p.u.]	[p.u.]	[hr]	[hr]
coal_lig_st	2.75	60000	0.3	0.50	2602.36	3580.95	50	0	0.006	0.006	0.05	24	12
coal_sub_st	2.75	60000	0.3	0.50	2602.36	3580.95	50	0	0.006	0.006	0.05	24	12
ng_cc	1.7	14000	1	0.30	589.47	9250.91	300	0	0.02	0.02	0.17	6	12
ng_gt	1	11000	6	0.10	187.72	7811.23	Inf	1	0.12	0.12	1	0	0
ng_st	1.7	14000	1	0.70	1172.48	1971.16	50	0	0.02	0.02	0.17	12	12
u235_st	1	90000	0.1	0.80	0	100000	1	0	0	0	0	0	24
wind	2	40000	1	0.00	0	0	Inf	0	0	0	0	0	0

F. Full IEEE RTS MIP Heuristics Results Table

						mip gap		Total Cost		CO2		energy mix	commitment		power		Time	
System	**# gens**	**# clust**	**# hrs**	**units**	**cheat**	**target**	**actual**	**$M**	**% diff**	**Mt**	**% diff**	**avg diff**	**# diff**	**avg |diff|**	**# diff**	**avg |diff|**	**sec**	**speed-up**
1 day, 1x scaling, Baseline=Optimal																		
IEEE RTS96	26	26	24	separate	0	0	0	0.72	NA	0.0	0.000%			Baseline			12.1	1
IEEE RTS96	26	26	24	perturbed	0	0	0	0.72	0.000%	0.0	0.000%	0.000%	0	0.000%	0	0.000%	7.0	1.7
IEEE RTS96	26	8	24	clustered	0	0	0	0.72	0.000%	0.0	0.000%	0.000%	0	0.000%	2	0.000%	0.2	63
IEEE RTS96	26	26	24	separate	0	0.10%	0.09%	0.72	0.000%	0.0	0.000%	0.000%	0	0.000%	0	0.000%	3.2	4
IEEE RTS96	26	26	24	perturbed	0	0.10%	0.10%	0.72	0.000%	0.0	0.000%	0.000%	0	0.000%	0	0.000%	2.1	6
IEEE RTS96	26	8	24	clustered	0	0.10%	0.10%	0.72	0.000%	0.0	0.000%	0.000%	0	0.000%	2	0.000%	0.2	69
1 day, 10x scaling, Baseline=Optimal																		
IEEE RTS96	260	260	24	separate	0	0	0.018%	6.74	NA	0.3	0.000%			Baseline			36000.7	1
IEEE RTS96	260	260	24	perturbed	0	0	0.016%	6.74	0.000%	0.3	-0.114%	0.022%	9	0.182%	18	0.030%	36000.7	1
IEEE RTS96	260	8	24	clustered	0	0	0	6.73	-0.233%	0.3	-0.458%	0.675%	49	1.136%	80	0.612%	5.6	6416
IEEE RTS96	260	260	24	separate	0.001	0	0	6.74	0.031%	0.3	0.012%	0.106%	37	0.215%	64	0.099%	517.4	70
IEEE RTS96	260	260	24	perturbed	0.001	0	0.105%	6.75	0.088%	0.3	-0.245%	0.852%	46	1.508%	80	0.750%	36000.7	1
IEEE RTS96	260	8	24	clustered	0.001	0	0	6.73	-0.208%	0.3	-0.459%	0.678%	45	1.137%	79	0.615%	0.2	167445
IEEE RTS96	260	260	24	separate	0	0.10%	0.09%	6.74	0.031%	0.3	0.012%	0.106%	37	0.215%	64	0.099%	254.2	142
IEEE RTS96	260	260	24	perturbed	0	0.10%	0.08%	6.74	0.037%	0.3	-0.223%	0.717%	39	1.042%	60	0.628%	338.2	106
IEEE RTS96	260	8	24	clustered	0	0.10%	0.05%	6.73	-0.208%	0.3	-0.459%	0.678%	45	1.137%	79	0.615%	0.2	174761
IEEE RTS96	260	260	24	separate	0.001	0.10%	0.00%	6.74	0.031%	0.3	0.012%	0.106%	37	0.215%	64	0.099%	546.1	66
IEEE RTS96	260	260	24	perturbed	0.001	0.10%	0.10%	6.75	0.088%	0.3	-0.245%	0.852%	46	1.508%	80	0.750%	36000.6	1
IEEE RTS96	260	8	24	clustered	0.001	0.10%	0.00%	6.73	-0.208%	0.3	-0.459%	0.678%	45	1.137%	79	0.615%	0.2	169017
1 day, 10x scaling, Baseline=0.1% MIP gap																		
IEEE RTS96	260	260	24	separate	0	0.10%	0.09%	6.74	NA	0.3	0.000%			Baseline			254.2	1
IEEE RTS96	260	260	24	perturbed	0	0.10%	0.08%	6.74	0.006%	0.3	-0.235%	0.616%	44	0.933%	71	0.542%	338.2	1
IEEE RTS96	260	8	24	clustered	0	0.10%	0.05%	6.73	-0.240%	0.3	-0.471%	0.579%	44	1.057%	77	0.531%	0.2	1234
IEEE RTS96	260	260	24	separate	0.001	0.10%	0.00%	6.74	0.000%	0.3	0.000%	0.000%	0	0.000%	0	0.000%	546.1	0.5
IEEE RTS96	260	260	24	perturbed	0.001	0.10%	0.10%	6.75	0.057%	0.3	-0.257%	0.749%	35	1.327%	60	0.651%	36000.6	0
IEEE RTS96	260	8	24	clustered	0.001	0.10%	0.00%	6.73	-0.240%	0.3	-0.471%	0.579%	44	1.057%	77	0.531%	0.2	1193
7 day, 1x scaling, Baseline=Optimal																		
IEEE RTS96	26	26	168	separate	0	0	0.044%	4.29	NA	0.2	0.000%			Baseline			36004.4	1
IEEE RTS96	26	26	168	perturbed	0	0	0.075%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	0	0.000%	36030.8	1
IEEE RTS96	26	26	168	priority list	0	0	0.000579	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	36001.3	1
IEEE RTS96	26	8	168	clustered	0	0	0	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	41.7	864
IEEE RTS96	26	26	168	separate	0.001	0	0.000%	4.30	0.055%	0.2	0.001%	0.022%	56	0.286%	124	0.035%	23589.9	2
IEEE RTS96	26	26	168	perturbed	0.001	0	0.001034	4.29	0.030%	0.2	0.046%	0.014%	18	0.082%	46	0.026%	36003.1	1
IEEE RTS96	26	26	168	priority list	0.001	0	0.00%	4.29	0.034%	0.2	0.017%	0.005%	9	0.035%	21	0.008%	1859.9	19
IEEE RTS96	26	8	168	clustered	0.001	0	0.00%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	27.6	1306
IEEE RTS96	26	26	168	separate	0	0.10%	0.10%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	0	0.000%	2543.7	14
IEEE RTS96	26	26	168	perturbed	0	0.10%	0.10%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	0	0.000%	11029.7	3
IEEE RTS96	26	26	168	priority list	0	0.10%	0.10%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	4892.5	7
IEEE RTS96	26	8	168	clustered	0	0.10%	0.10%	4.29	0.023%	0.2	0.015%	0.014%	7	0.024%	17	0.016%	9.3	3891
IEEE RTS96	26	26	168	separate	0.001	0.10%	0.00%	4.30	0.055%	0.2	0.001%	0.022%	56	0.286%	124	0.035%	19717.8	2
IEEE RTS96	26	26	168	perturbed	0.001	0.10%	0.00%	4.29	0.030%	0.2	0.046%	0.014%	18	0.082%	46	0.026%	33672.2	1
IEEE RTS96	26	26	168	priority list	0.001	0.10%	0.00%	4.29	0.034%	0.2	0.017%	0.005%	9	0.035%	21	0.008%	1918.2	19
IEEE RTS96	26	8	168	clustered	0.001	0.10%	0.00%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	23.7	1519
7 day, 1x scaling, Baseline=0.1% MIP gap																		
IEEE RTS96	26	260	168	separate	0	0.10%	0.10%	4.29	NA	0.2	0.000%			Baseline			2543.7	1
IEEE RTS96	26	260	168	perturbed	0	0.10%	0.10%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	0	0.000%	11029.7	0
IEEE RTS96	26	8	168	priority list	0	0.10%	0.10%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	4892.5	1
IEEE RTS96	26	260	168	clustered	0	0.10%	0.10%	4.29	0.023%	0.2	0.015%	0.014%	7	0.024%	17	0.016%	9.3	274.9
IEEE RTS96	26	8	168	separate	0.001	0.10%	0.00%	4.30	0.055%	0.2	0.001%	0.022%	56	0.286%	124	0.035%	19717.8	0
IEEE RTS96	26	260	168	perturbed	0.001	0.10%	0.00%	4.29	0.030%	0.2	0.046%	0.014%	18	0.082%	46	0.026%	33672.2	0.1
IEEE RTS96	26	260	168	priority list	0.001	0.10%	0.00%	4.29	0.034%	0.2	0.017%	0.005%	9	0.035%	21	0.008%	1918.2	1
IEEE RTS96	26	8	168	clustered	0.001	0.10%	0.00%	4.29	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	23.7	107
7 day, 3x scaling, Baseline=Optimal																		
IEEE RTS96	78	78	168	separate	0	0	0.085%	12.16	NA	0.6	0.000%			Baseline			36032.8	1
IEEE RTS96	78	78	168	perturbed	0	0	0.086%	12.16	0.002%	0.6	0.064%	0.035%	123	0.334%	247	0.082%	36028.4	1
IEEE RTS96	78	78	168	priority list	0	0	0.002924	12.19	0.208%	0.6	-0.021%	0.060%	211	0.605%	396	0.155%	36012.4	1
IEEE RTS96	78	8	168	clustered	0	0	0.000195	12.16	-0.028%	0.6	0.045%	0.026%	141	0.416%	318	0.075%	36001.1	1
IEEE RTS96	78	78	168	separate	0.001	0	0.176%	12.17	0.091%	0.6	0.125%	0.080%	129	0.736%	278	0.166%	36002.2	1
IEEE RTS96	78	78	168	perturbed	0.001	0	0.001737	12.17	0.088%	0.6	0.133%	0.091%	172	0.516%	323	0.156%	36035.0	1
IEEE RTS96	78	78	168	priority list	0.001	0.00%	0.25%	12.18	0.166%	0.6	0.041%	0.067%	218	0.643%	410	0.138%	36012.4	1
IEEE RTS96	78	8	168	clustered	0.001	0.00%	0.00%	12.16	-0.027%	0.6	0.017%	0.022%	98	0.279%	238	0.063%	6426.4	6
IEEE RTS96	78	78	168	separate	0	0.10%	0.10%	12.16	0.010%	0.6	0.053%	0.018%	102	0.329%	228	0.052%	5980.8	6
IEEE RTS96	78	78	168	perturbed	0	0.10%	0.09%	12.16	0.007%	0.6	0.030%	0.008%	85	0.327%	197	0.038%	23367.1	2
IEEE RTS96	78	78	168	priority list	0	0.10%	0.29%	12.19	0.208%	0.6	-0.021%	0.060%	211	0.605%	396	0.155%	36012.4	1
IEEE RTS96	78	8	168	clustered	0	0.10%	0.10%	12.17	0.039%	0.6	-0.036%	0.053%	172	0.462%	341	0.129%	42.0	857
IEEE RTS96	78	78	168	separate	0.001	0.10%	0.18%	12.17	0.091%	0.6	0.125%	0.080%	129	0.736%	278	0.166%	36001.8	1
IEEE RTS96	78	78	168	perturbed	0.001	0.10%	0.17%	12.17	0.088%	0.6	0.133%	0.091%	172	0.516%	323	0.156%	36011.7	1
IEEE RTS96	78	78	168	priority list	0.001	0.10%	0.25%	12.18	0.166%	0.6	0.041%	0.067%	218	0.643%	410	0.138%	36012.4	1
IEEE RTS96	78	8	168	clustered	0.001	0.10%	0.00%	12.16	-0.027%	0.6	0.017%	0.022%	98	0.279%	238	0.063%	4912.1	7
7 day, 3x scaling, Baseline=0.1% MIP gap																		
IEEE RTS96	78	260	168	separate	0	0.10%	0.10%	12.16	NA	0.6	0.000%			Baseline			5980.8	1
IEEE RTS96	78	260	168	perturbed	0	0.10%	0.09%	12.16	-0.002%	0.6	-0.023%	0.011%	93	0.397%	188	0.048%	23367.1	0
IEEE RTS96	78	8	168	priority list	0	0.10%	0.29%	12.19	0.199%	0.6	-0.074%	0.065%	192	0.602%	358	0.148%	36012.4	0
IEEE RTS96	78	260	168	clustered	0	0.10%	0.10%	12.17	0.029%	0.6	-0.089%	0.060%	173	0.516%	319	0.130%	42.0	142.3
IEEE RTS96	78	8	168	separate	0.001	0.10%	0.18%	12.17	0.082%	0.6	0.072%	0.074%	135	0.735%	275	0.175%	36001.8	0
IEEE RTS96	78	260	168	perturbed	0.001	0.10%	0.17%	12.17	0.079%	0.6	0.080%	0.074%	168	0.509%	316	0.143%	36011.7	0.2
IEEE RTS96	78	260	168	priority list	0.001	0.10%	0.25%	12.18	0.156%	0.6	-0.012%	0.084%	197	0.609%	367	0.119%	36012.4	0
IEEE RTS96	78	8	168	clustered	0.001	0.10%	0.00%	12.16	-0.037%	0.6	-0.035%	0.034%	95	0.280%	202	0.055%	4912.1	1

G. IEEE RTS Unit Commitment Simplifications Results

					Simplifications						mip gap		Total Cost		CO2		energy mix	commitment		power		Time		
System	# gens	# clust	# hrs	units	ignore min up & down times	max_start	UC integer threshold (MW)	combine rsrv	afine fuel use	cheat	target	actual	$M	% diff	Mt	% diff	avg diff	# diff	avg |diff|	# diff	avg |diff|	sec	speed-up from baseline	additional speedup for clustering
7 day, 1x scaling, Baseline=0.1% MIP gap, No Cheat																								
IEEE RTS96	26	26	168	separate	-	-	NaN	-	-	0	0.10%	0.10%	4	NA	0.2	0.000%	Baseline					2543.7	1.0	
IEEE RTS96	26	8	168	clustered	-	-	NaN	-	-	0	0.10%	0.10%	4	0.023%	0.2	0.015%	0.014%	7	0.024%	17	0.016%	9.3	274.9	275
IEEE RTS96	26	26	168	separate	-	-	5	-	-	0	0.10%	0.10%	4	-0.052%	0.2	-0.049%	0.015%	75	0.328%	185	0.045%	362.1	7.0	
IEEE RTS96	26	8	168	clustered	-	-	5	-	-	0	0.10%	0.10%	4	-0.036%	0.2	0.002%	0.035%	60	0.243%	161	0.043%	7.8	324.1	46
IEEE RTS96	26	26	168	separate	-	-	20	-	-	0	0.10%	0.10%	4	-0.127%	0.2	-0.087%	0.029%	69	0.274%	271	0.073%	153.6	16.6	
IEEE RTS96	26	8	168	clustered	-	-	20	-	-	0	0.10%	0.09%	4	-0.105%	0.2	-0.208%	0.025%	80	0.315%	290	0.083%	1.4	1875.9	113
IEEE RTS96	26	26	168	separate	-	-	0	-	Y	0	0.10%	0.09%	4	-0.212%	0.2	-0.154%	0.084%	0	0.000%	55	0.105%	532.2	4.8	
IEEE RTS96	26	8	168	clustered	-	-	0	-	Y	0	0.10%	0.04%	4	-0.195%	0.2	-0.140%	0.089%	5	0.018%	69	0.118%	27.6	92.0	19
IEEE RTS96	26	26	168	separate	-	-	0	Y	-	0	0.10%	0.10%	4	-0.194%	0.2	-0.035%	0.061%	126	1.458%	407	0.155%	560.3	4.5	
IEEE RTS96	26	8	168	clustered	-	-	0	Y	-	0	0.10%	0.08%	4	-0.153%	0.2	-0.083%	0.054%	120	1.279%	397	0.152%	27.0	94.3	21
IEEE RTS96	26	26	168	separate	Y	-	0	-	-	0	0.10%	0.10%	4	-0.652%	0.2	0.043%	0.170%	89	0.754%	176	0.231%	1191.0	2.1	
IEEE RTS96	26	8	168	clustered	Y	-	0	-	-	0	0.10%	0.06%	4	-0.630%	0.2	0.046%	0.166%	95	0.775%	189	0.245%	4.6	558.7	262
IEEE RTS96	26	26	168	separate	Y	-	0	-	-	0	0.10%	0.10%	4	-0.645%	0.2	0.055%	0.172%	93	0.776%	188	0.236%	236.0	10.8	
IEEE RTS96	26	8	168	clustered	Y	-	0	-	-	0	0.10%	0.09%	4	-0.599%	0.2	0.066%	0.164%	96	0.791%	190	0.249%	8.4	303.7	28
IEEE RTS96	26	26	168	separate	-	Y	0	-	-	0	0.10%	0.10%	4	0.000%	0.2	0.000%	0.000%	0	0.000%	1	0.000%	3916.3	0.6	
IEEE RTS96	26	8	168	clustered	-	Y	0	-	-	0	0.10%	0.10%	4	0.062%	0.2	-0.073%	0.015%	44	0.283%	77	0.033%	19.9	128.0	197
7 day, 3x scaling, Baseline=0.1% MIP gap, No Cheat																								
IEEE RTS96	78	78	168	separate	-	-	0	-	-	0	0.10%	0.10%	12	NA	0.6	0.000%	Baseline					5980.8	1.0	
IEEE RTS96	78	8	168	clustered	-	-	0	-	-	0	0.10%	0.10%	12	0.029%	0.6	-0.089%	0.060%	173	0.516%	319	0.130%	42.0	142.3	142
IEEE RTS96	78	78	168	separate	-	-	5	-	-	0	0.10%	0.08%	12	-0.015%	0.6	-0.044%	0.030%	141	0.584%	353	0.061%	105.9	56.5	
IEEE RTS96	78	8	168	clustered	-	-	5	-	-	0	0.10%	0.09%	12	0.006%	0.6	-0.017%	0.025%	109	0.337%	304	0.056%	1.3	4562.0	81
IEEE RTS96	78	78	168	separate	-	-	20	-	-	0	0.10%	0.08%	12	-0.013%	0.6	-0.046%	0.028%	158	0.639%	419	0.077%	113.7	52.6	
IEEE RTS96	78	8	168	clustered	-	-	20	-	-	0	0.10%	0.08%	12	-0.014%	0.6	-0.038%	0.035%	109	0.391%	348	0.056%	1.4	4375.1	83
IEEE RTS96	78	78	168	separate	-	-	0	-	Y	0	0.10%	0.08%	12	-0.143%	0.6	-0.144%	0.386%	216	0.801%	456	0.435%	1619.0	3.7	
IEEE RTS96	78	8	168	clustered	-	-	0	-	Y	0	0.10%	0.10%	12	-0.107%	0.6	-0.174%	0.360%	199	0.597%	456	0.411%	22.9	260.8	71
IEEE RTS96	78	78	168	separate	-	-	0	Y	-	0	0.10%	0.10%	12	-0.052%	0.6	0.004%	0.065%	171	0.488%	405	0.143%	5197.3	1.2	
IEEE RTS96	78	8	168	clustered	-	-	0	Y	-	0	0.10%	0.07%	12	-0.071%	0.6	0.050%	0.054%	160	0.515%	404	0.116%	2.9	2097.1	1822
IEEE RTS96	78	78	168	separate	Y	-	0	-	-	0	0.10%	0.16%	12	-0.044%	0.6	-0.148%	0.150%	223	0.831%	418	0.225%	36064.5	0.2	
IEEE RTS96	78	8	168	clustered	Y	-	0	-	-	0	0.10%	0.08%	12	-0.108%	0.6	-0.092%	0.201%	189	0.678%	372	0.212%	12.0	498.1	3003
IEEE RTS96	78	78	168	separate	Y	-	0	-	-	0	0.10%	0.08%	12	-0.133%	0.6	-0.086%	0.185%	219	0.732%	413	0.238%	27739.8	0.2	
IEEE RTS96	78	8	168	clustered	Y	-	0	-	-	0	0.10%	0.10%	12	-0.103%	0.6	-0.058%	0.171%	175	0.621%	333	0.173%	2.7	2247.6	10425
IEEE RTS96	78	78	168	separate	-	Y	0	-	-	0	0.10%	0.08%	12	-0.015%	0.6	-0.043%	0.030%	182	0.645%	327	0.120%	2962.9	2.0	
IEEE RTS96	78	8	168	clustered	-	Y	0	-	-	0	0.10%	0.09%	12	0.003%	0.6	-0.055%	0.040%	141	0.394%	296	0.085%	7.9	759.3	376

H. ERCOT 2007 Unit Commitment Simplifications Results

					Simplifications					mip gap		Total Cost		CO2		energy mix	commitment		power		Time		
System	# gens	# clust	# hrs	units	ignore min up & down times	max_start	UC integer threshold (MW)	combine rsrv	afine fuel use	target	actual	$M	% diff	Mt	% diff	avg diff	# diff	avg |diff|	# diff	avg |diff|	sec	speed-up from baseline	additional speedup for clustering
1 week,No Cheat, Baseline=Full Problem																							
ERCOT2007	205	205	168	separate	-	-	0	-	Y	0.10%	0.06%	349	NA	3.27	0.000%	Baseline					4517.2	1.0	
ERCOT2007	205	7	168	clustered	-	-	0	-	Y	0.10%	0.10%	352	0.785%	3.31	1.253%	0.377%	336	1.150%	625	0.533%	2.2	2073.1	2073
ERCOT2007	205	205	168	separate	-	-	5	-	Y	0.10%	0.06%	349	0.000%	3.27	0.000%	0.000%	0	0.000%	0	0.000%	3693.8	1.2	
ERCOT2007	205	7	168	clustered	-	-	5	-	Y	0.10%	0.10%	352	0.785%	3.31	1.253%	0.377%	336	1.150%	625	0.533%	1.9	2409.2	1970
ERCOT2007	205	205	168	separate	-	-	20	-	Y	0.10%	0.03%	349	-0.033%	3.28	0.103%	0.051%	298	1.071%	508	0.076%	951.7	4.7	
ERCOT2007	205	7	168	clustered	-	-	20	-	Y	0.10%	0.01%	352	0.696%	3.31	1.174%	0.330%	233	0.971%	598	0.504%	0.5	9798.7	2064
ERCOT2007	205	205	168	separate	-	-	0	Y	Y	0.10%	0.01%	349	-0.026%	3.28	0.194%	0.141%	296	2.293%	546	0.187%	376.6	12.0	
ERCOT2007	205	7	168	clustered	-	-	0	Y	Y	0.10%	0.03%	351	0.555%	3.31	1.293%	0.340%	140	0.665%	598	0.524%	0.3	16486.1	1374
ERCOT2007	205	205	168	separate	Y	Y	0	-	Y	0.10%	0.03%	349	-0.138%	3.28	0.316%	0.090%	146	0.449%	492	0.141%	508.0	8.9	
ERCOT2007	205	7	168	clustered	Y	Y	0	-	Y	0.10%	0.04%	352	0.728%	3.31	1.203%	0.334%	318	1.169%	601	0.504%	1.5	3005.5	338
ERCOT2007	205	205	168	separate	-	-	0	-	Y	0.10%	0.05%	349	-0.009%	3.27	0.039%	0.013%	266	0.418%	532	0.057%	1681.9	2.7	
ERCOT2007	205	7	168	clustered	-	-	0	-	Y	0.10%	0.05%	352	0.739%	3.31	1.210%	0.346%	307	1.114%	613	0.510%	1.5	2993.5	1115
13 weeks, Combined Reserves, No Cheat, Baseline=Full Problem, Timeout Extended to 60hrs. Run using faster machine, compare with care																							
ERCOT2007	205	205	2184	separate	-	-	0	Y	Y	0.10%	0.06%	6147	NA	55.62	0.000%	Baseline					37633.5	1.0	
ERCOT2007	205	7	2184	clustered	-	-	0	Y	Y	0.10%	0.00%	6203	0.907%	56.04	0.744%	0.243%	5064	2.627%	7354	0.412%	7.2	5234.1	1115
ERCOT2007	205	205	2184	separate	-	-	5	Y	Y	0.10%	0.06%	6147	0.000%	55.62	0.000%	0.000%	0	0.000%	0	0.000%	35203.2	1.1	
ERCOT2007	205	7	2184	clustered	-	-	5	Y	Y	0.10%	0.00%	6203	0.907%	56.04	0.744%	0.243%	5064	2.627%	7354	0.412%	6.6	5704.6	1116
ERCOT2007	205	205	2184	separate	-	-	20	Y	Y	0.10%	0.00%	6142	-0.082%	55.60	-0.039%	0.015%	3506	0.675%	4921	0.083%	10630.6	3.5	
ERCOT2007	205	7	2184	clustered	-	-	20	Y	Y	0.10%	0.00%	6203	0.902%	56.03	0.738%	0.246%	5046	2.600%	7330	0.413%	6.1	6211.2	1117
ERCOT2007	205	205	2184	separate	Y	Y	0	Y	Y	0.10%	0.41%	6166	0.300%	55.76	0.248%	0.055%	5008	0.812%	5957	0.489%	216044.2	0.2	
ERCOT2007	205	7	2184	clustered	Y	Y	0	Y	Y	0.10%	0.00%	6186	0.623%	55.91	0.512%	0.474%	5071	2.693%	7320	0.587%	18.0	2087.4	1118
ERCOT2007	205	205	2184	separate	-	-	0	Y	Y	0.10%	0.02%	6158	0.170%	55.73	0.191%	0.116%	4283	1.257%	5164	0.302%	72667.3	0.5	
ERCOT2007	205	7	2184	clustered	-	-	0	Y	Y	0.10%	0.00%	6203	0.907%	56.04	0.744%	0.242%	5069	2.632%	7357	0.418%	7.4	5076.7	1119

I. ERCOT 2007 Cluster Comparison

					mip gap		Total Cost		CO2		energy mix	commitment		power		Time	
System	# gens	# clust	# hrs	units	target	actual	$M	% diff	Mt	% diff	avg diff	# diff	avg |diff|	# diff	avg |diff|	sec	speed-up
13 weeks, Combined Reserves, Afine Fuel Use, No Cheat, Baseline=Full Problem, Timeout Extended to 60hrs.																	
ERCOT2007	205	205	2184	separate	0.10%	0.06%	6147	NA	55.6	0.000%	Baseline					37633.5	1
ERCOT2007	205	90	2184	clust_plant	0.10%	0.07%	6133	-0.232%	55.6	0.008%	0.082%	3869	0.842%	7311	0.117%	7701.8	5
ERCOT2007	205	17	2184	clust_age	0.10%	0.02%	6468	5.211%	56.0	0.601%	1.753%	9601	2.633%	12041	1.970%	33.1	1135
ERCOT2007	205	17	2184	clust_size	0.10%	0.03%	6192	0.727%	56.0	0.667%	0.222%	5072	2.454%	5222	0.319%	33.4	1128
ERCOT2007	205	17	2184	clust_efficiency	0.10%	0.01%	6153	0.100%	55.7	0.062%	0.136%	4436	0.898%	5213	0.208%	41.9	898
ERCOT2007	205	7	2184	clustered	0.10%	0.00%	6203	0.907%	56.0	0.744%	0.243%	5064	2.627%	7354	0.412%	7.2	5234
1 year, Combined Reserves, Afine Fuel, No Cheat, Baseline=Full Problem, Run on standard machine with standard 10hr timelimit																	
ERCOT2007	205	205	8760	separate	0.10%	No Solution					No solution for Comparison					10hr Timeout	
ERCOT2007	205	90	8760	clust_plant	0.10%	No Solution										10hr Timeout	
ERCOT2007	205	17	8760	clust_age	0.10%	0.10%	22759	-	200.0							34230.6	-
ERCOT2007	205	17	8760	clust_size	0.10%	0.07%	21881	-	201.3							10722.1	-
ERCOT2007	205	17	8760	clust_efficiency	0.10%	0.13%	21781	-	200.2							10hr Timeout	
ERCOT2007	205	7	8760	clustered	0.10%	0.01%	21933	-	201.7							129.5	-

J. ERCOT Scaling Comparison

					mip gap		Problem size (Before CPLEX pre-solve)				Time	
System	# gens	# clust	# hrs	units	target	actual	# equations	# variables	# discretes	# non-zeros	sec	speed-up
No Cheat, 0.1% MIP Gap, Separate Reserves, All runs on standard machines with 10hour timeout												
ERCOT2007	205	205	24	separate	0.10%	0.00%	63,786	50,008	4,896	293,093	6	1.0
ERCOT2007	205	90	24	clust_plant	0.10%	0.01%	28,290	21,592	2,136	133,589	4	1.6
ERCOT2007	205	17	24	clust_size	0.10%	0.02%	5,394	3,928	384	25,997	0	37.5
ERCOT2007	205	7	24	clustered	0.10%	0.00%	2,130	1,480	144	10,109	0	71.7
ERCOT2007	205	205	48	separate	0.10%	0.03%	132,162	97,504	9,792	654,317	30	1.0
ERCOT2007	205	90	48	clust_plant	0.10%	0.10%	59,058	42,592	4,272	298,445	7	4.0
ERCOT2007	205	17	48	clust_size	0.10%	0.04%	10,914	7,696	768	55,565	0	67.1
ERCOT2007	205	7	48	clustered	0.10%	0.018%	4,290	2,896	288	22,157	0	314.2
ERCOT2007	205	205	72	separate	0.10%	0.06%	198,234	146,248	14,688	981,461	104	1.0
ERCOT2007	205	90	72	clust_plant	0.10%	0.07%	88,578	63,880	6,408	447,653	42	2.5
ERCOT2007	205	17	72	clust_size	0.10%	0.02%	16,362	11,536	1,152	83,333	1	157.2
ERCOT2007	205	7	72	clustered	0.10%	0.024%	6,426	4,336	432	33,221	0	695.0
ERCOT2007	205	205	96	separate	0.10%	0.03%	264,306	194,992	19,584	1,308,605	140	1.0
ERCOT2007	205	90	96	clust_plant	0.10%	0.08%	118,098	85,168	8,544	596,861	22	6.3
ERCOT2007	205	17	96	clust_size	0.10%	0.03%	21,810	15,376	1,536	111,101	1	99.2
ERCOT2007	205	7	96	clustered	0.10%	0.020%	8,562	5,776	576	44,285	0	758.9
ERCOT2007	205	205	168	separate	0.10%	0.06%	446,394	349,960	34,272	2,068,949	4,517	1.0
ERCOT2007	205	90	168	clust_plant	0.10%	0.06%	197,922	151,048	14,952	943,685	435	10.4
ERCOT2007	205	17	168	clust_size	0.10%	0.04%	37,650	27,400	2,688	186,173	10	442.7
ERCOT2007	205	7	168	clustered	0.10%	0.096%	14,802	10,264	1,008	74,957	2	2073.1
ERCOT2007	205	205	2184	separate	0.10%	NA	5,802,906	4,549,288	445,536	26,895,989	Timeout	
ERCOT2007	205	90	2184	clust_plant	0.10%	0.102%[A]	2,572,770	1,963,432	194,376	12,267,557	36030.6[A]	1.0
ERCOT2007	205	17	2184	clust_size	0.10%	0.099%	489,234	356,008	34,944	2,419,901	897	3.4
ERCOT2007	205	7	2184	clustered	0.10%	0.01%	192,210	133,240	13,104	974,093	48	278.0
ERCOT2007	205	205	8760	separate	0.10%	NA	23,275,338	18,247,096	1,787,040	107,879,430	Timeout	
ERCOT2007	205	90	8760	clust_plant	0.10%	NA	10,319,298	7,875,256	779,640	49,204,949	Timeout	
ERCOT2007	205	17	8760	clust_size	0.10%	0.07%	1,962,258	1,427,896	140,160	9,706,109	10,722	1.0
ERCOT2007	205	7	8760	clustered	0.10%	0.01%	770,898	534,376	52,560	3,906,989	129	82.8

[A] 13wk (2184 hr) cluster by plant results included despite time out at 10hrs, because MIP gap was only 0.002% greater than target