Validating Unit Commitment Models: A Case for Benchmark Test Systems

Alexander C. Melhorn and Mingsong Li
School of Electrical and Electronic Engineering
University College Dublin
Dublin, Ireland
Email: acmelhorn@gmail.com

Paula Carroll
School of Business
University College Dublin
Dublin, Ireland
Email: paula.carroll@ucd.ie

Damian Flynn, Senior Member IEEE
School of Electrical and Electronic Engineering
University College Dublin
Dublin, Ireland
Email: damian.flynn@ucd.ie

Abstract—Due to increasing penetration of non-traditional power system resources; e.g. renewable generation, electric vehicles, demand response, etc. and computational power there has been an increased interest in research on unit commitment. It therefore may be important to take another look at how unit commitment models and algorithms are validated especially as improvements in solutions and algorithmic performance are desired to combat the added complexity of additional constraints. This paper explores an overview of the current state of unit commitment models and algorithms, and finds improvements for both comparing and validating models with benchmark test systems. Examples are provided discussing the importance for a standard benchmark test system(s) and why it is needed to compare and validate the real world performance of unit commitment models.

Index Terms—Unit Commitment (UC), MILP, MIP, Benchmark test systems

I. INTRODUCTION

The unit commitment (UC) problem and associated economic dispatch (ED) problem have been studied widely in the engineering literature, but have received less attention in the combinatorial optimization literature. From 1966 to 2005, there were over 650 IEEE articles on UC research. Since 2006 there have been over 875 articles. Less than 100 articles in the Operational Research and Management Science (ORMS) literature were found over the same period of time.

The UC problem can be stated as that of determining the combination of electricity generating units to commit or decommit, so as to meet the expected power demand and satisfy the operating constraints of the generating units and of the system as a whole. The UC problem is linked to the ED problem which seeks to determine the power production of committed units. The interlinked problems are solved so as to minimize costs and/or other factors.

In recent times, these problems have been the subject of renewed interest due to competitive market systems, a focus on integrating renewable energy sources, and security and stability issues arising from the interconnection of national and regional systems. These concerns along with improvements in computational power have driven interest from the community of practitioners for improved UC algorithmic performance.

In practice, transmission system operators (TSOs) solve UC mixed integer (linear) programming (MILP) models on a day-ahead basis [1]. These UC models may contain hundreds of generators, sets of generator constraints, estimated hourly demand data and renewable generation profiles, and additional operational constraints. Subsequent in-day schedules may determine modifications to the day-ahead schedule in real time based on actual demand, generation capacity and unforeseen events. TSOs have indicated a need to optimize models at 15 minute demand intervals rather than hourly, to increase the ability of the electricity grid to respond to fluctuating availability of renewable power sources and demand reserves.

A key concern when proposing new UC models is their empirical performance on real world data sets; i.e. feasibility (both in finding an optimal solution and in real world operational procedures), optimal results, and solution speed; especially compared to previous methods. The most cited UC algorithms and associated test data are summarized within this paper [2]–[14]. In many cases computational results are given for small theoretical test instances, or based on a UC instance in [2] which dates from 1996. Much of the reported test data is incomplete so results cannot be easily replicated. Furthermore, a valid comparison is not always possible if it is not clear which constraints have been applied. It is difficult to evaluate proposed UC solution algorithms in the absence of real world test data in a standard format. Without proper validation, UC models can be unknowingly providing sub-optimal results, be infeasible for a variety of runtime instances (constraints may not even be bound in certain instances) or infeasible in real world operations, providing a false sense of the solution time, etc. which can increase the need for utilizing reserves, close to real-time interventions, and reduce TSO profits. A small percentage change in the UC solution was predicted to save PJM over 150 million dollars annually and a similar change was predicted to save CAISO millions annually [1].

IEEE benchmark test systems are available for various other areas of power system research, including, but not limited to, system reliability [15], dynamic power flow analysis [16], and distribution systems [17]–[19]. Standardized benchmark instances are also available in other ORMS areas [20], [21]. Some recent work has been conducted for UC test systems [22], [23].

An overview of the UC problem and a basic MILP UC model is presented in Sec. II. The most cited UC algorithms
and their associated test data are summarized in Sec. III. Sec. IV discusses the case for and various methods of proper UC validation/benchmarking. The paper concludes with Sec. V summarizing the paper and providing some final thoughts on validating UC models.

II. THE UNIT COMMITMENT PROBLEM

TSOs typically rely on MILP approaches compared to the traditional Lagrangian relaxation (LR) due to its ability to find the global optimal solution and the increase in computational power available [1]. A MILP implementation of a basic UC model, following [3], [24], is summarized below.

A. Objective

Typically, the objective of UC is to minimize operating costs of the system over a planning horizon, usually a (rolling) daily horizon at an hourly time-step. UC models traditionally minimize power production, start-up and shut-down costs of each generator are included. Due to changing policies, an increase of renewable generation, energy storage, demand response and smart grid initiatives, other costs may be included. Work is ongoing on how to best capture the cost of renewable sources such as wind, reserve, cycling and emissions in the objective function:

\[
\min \sum_{k \in K} \sum_{g \in G} c^p_{g,k} + c^u_{g,k} + c^d_{g,k},
\]

where \( g \in G \) is the index of generating units, \( k \in K \) is the index of time-steps, and \( c^p_{g,k}, c^u_{g,k}, c^d_{g,k} \) are the power production, start-up and shut-down costs respectively.

B. Basic Constraints

The primary constraint is the production constraint (2). Total production of the units at a given time must equal the load demand. An approach to integrating renewable sources, such as wind, into the grid, is to reduce the demand by the amount of renewable capacity available to give net demand. This approach, however, does not adequately reflect the cost and impact of wind. The reserve constraint (3) ensures that the maximum production available is greater than or equal to the demand plus a given reserve target. Reserve can be classed as momentary, fast (spinning) or slow depending on which party is responsible for reacting to an outage [12]. For this formulation, the reserve does not influence the operational cost directly in (3). In practice TSOs may enforce multiple reserve constraints available for response at different time frames. In the event of a fault or other contingency on the system, this reserve power may be called upon. A number of reserve constraints may be specified, focusing on the system response over different time horizons. (Blackouts in recent years have increased the focus on the design and operation of secure grids.)

\[
\sum_{g \in G} P_{g,k} = D_k \quad \forall k \in K,
\]

\[
\sum_{g \in G} P_{g,k} \geq D_k + R_k \quad \forall k \in K.
\]

Each generator has a basic set of operating characteristics specified by the following operating parameters:

- \( P \) – Maximum output (MW)
- \( P \) – Minimum output (MW)
- \( UT \) – Minimum time a unit must stay online (up) once it has been switched online
- \( DT \) – Minimum time a unit must stay offline (down) once it has been switched offline
- \( IS \) – Initial State, no. time periods unit has been on (off) line at \( k = 0 \)
- \( a, b, c \) – Coefficients of quadratic power production cost function

The MILP implementation can be modeled as:

\[
P_g \cdot v_{g,k} \leq p_{g,k} \leq P_g \cdot v_{g,k} \quad \forall g \in G, k \in K
\]

\[
\underline{P}_g \cdot v_{g,k} \leq P_{g,k} \leq \bar{P}_g \cdot v_{g,k} \quad \forall g \in G, k \in K
\]

\[
c^u_{g,k} + c^d_{g,k} \leq c^p_{g,k} \quad \forall g \in G, k \in K
\]

where \( p_{g,k} \) and \( P_{g,k} \) are the power output and power availability, and \( v_{g,k} \) are binary variables set to 1 if unit \( g \) is committed (on) in time \( k \).

C. Additional Constraints

There are many other generator parameters that may be considered in addition to the set above. They include, but are not limited to:

- \( c^u \) – start-up cost
- \( c^d \) – shut-down cost
- \( RU, RD \) – ramp-up/down limits; rate of power change per time-step when a unit is running
- \( SU, SD \) – start-up/shut-down limits; rate of power change per time-step at start up/shut down

As systems vary so do their requirements for daily operation. Other constraints may be appropriate for the system in addition to the generation constraints. Again, these may include, but are not limited to:

- hydro generation
- \( P_0 \) – power output at start of planning horizon
- final state requirements
- operational requirements
- multiple reserve requirements
- maintenance schedules
- network constraints
- cycling constraints
- energy storage
- system stability limits
- fuel usage constraints
- plant crew considerations
- emission constraints (CO2, NOx, and SOx)
- system stability limits
In general the UC problem can be stated as determining the minimum cost of the generation dispatch schedule that meets an estimated demand and satisfies the system’s operating constraints. The basic problem is described in [24]. There is a large number of combinations of generators that could be committed in each time-step and complete enumeration on real instances is infeasible. An $NP$-hardness and $NP$-complete proof is given in [25] signifying the UC problem is very complex and difficult to solve, especially as the problem size grows.

An early summary of UC literature is given in [26]. The authors summarize a number of UC solution approaches such as priority lists, dynamic programming (DP), LR, MILP, simulated annealing (SA), expert systems and artificial neural networks. No test data or results are given in [26].

A more up-to-date review is given in [27]. It is noted that power from traditional thermal sources can be displaced by cheaper power from renewable sources, leading to more challenging UC problem instances due to increased variability and uncertainty from renewable generations stochastic nature. Solution approaches from the engineering literature are summarized in [27] and the list extended to include fuzzy systems, genetic algorithms (GA), evolutionary programming (EP), ant colony heuristics and hybrid approaches. Again no test data or results are given.

Table I gives a summary of the most widely cited articles concerning UC solution approaches. The reference, year of publication, number of citations, the type of approach employed and the problem size tested are all listed.

A GA is described in [2]. A 10 unit benchmark data set is provided, but there is no detail of the source of costs. No ramping data or shut down costs are given, although they are mentioned as part of the GA fitness function. The reserve requirement is assumed to be 10% of demand. Computational results are given for the 10 unit test data and larger test instances based on duplicating the 10 base units and scaling demand accordingly. A LR algorithm is used to find a near optimal solution against which the GA results are measured.

LR algorithms are also described in [6] along with computational results for a 100 unit system. The test system is summarized within the paper but full details are not made available. A 172 unit French system is tested in [10] using LR, but unfortunately the test data are not made available.

A MILP model is described in [11]. The branch-and-bound search space is reduced by making certain assumptions about cycling of units in the planning horizon. Cycling issues include the turning on and off of units, and ramping units at their maximum ramp rates which leads to increased wear-and-tear. No data or computational results are given.

A more detailed MILP formulation is given in [3]. Computational results for the MILP model are presented using the test data from [2]. Setting a 0.5% optimality gap, an integer solution is reported which is an improvement on the results given in [2], [5]. A common approach is used, similar to [2], whereby shut down costs are ignored. Ramping constraints are described but no ramp limits are given. Reserve is assumed to be 10% of demand.

A number of heuristic approaches have been used with some success. Evolutionary programming is employed in [5]. A Particle Swarm Optimization (PSO) algorithm is described in [14]. Computational results for the data from [2] are given for both these heuristics.

A number of works take the stochastic nature of the UC problem into account, i.e. the uncertainty in future demand profile and/or renewable generation [4], [7], [12]. The determination of reserve requirements based on their probabilistic nature is the focus in [12]. The impact of interconnect contracts are also addressed. Computational results for a small five unit test case are given.

A scenario tree is used in [7] to simulate the uncertainty of future demand and [4] uses a rolling planning method with scenario trees to model future demand and wind production scenarios. Each scenario has an intrinsic probability and the whole tree depicts the entire sample space. A number of deterministic sub-problems are solved together rather than a single estimated demand case. A more comprehensive solution is achieved compared with a conventional deterministic model. The authors note that high efficiency methods for solving the basic deterministic problem are needed, which remains the case today.

Some more recent work has focused on selecting strong MILP formulations [1], [28]–[33]. An analysis of valid inequalities is given in [1], [28], [29], with a focus on the start-up and shut-down ramp rates in [30] and combined cycle gas turbines in [31]. The min-up/down polytope associated with UC models is characterized in [32] with focus on the start-up costs in [33].

There are a number of other relevant articles where data are made available. A decomposition algorithm with test results on a 10 unit instance is given in [34]. Ramping rates are omitted and start-up costs are modeled as an exponential function of the number of time-steps that the unit has been off. Constraint programming is used in [35] to solve the 38 unit Taipower system. Ramp-up rates are given but initial states, shut-down costs and ramp-down rates are omitted.

A parallel GA is given in [36] which is tested on a 45 unit system. This paper offers the most complete test data set available during the exploration of the literature. Start-up cost coefficients for an exponential function are given. A start-up function is described in [24] that could be used to calculate hot and cold start costs. However, the initial conditions of the 45 unit system are such that the first 35 units must remain on for the duration of the time horizon, with only their power output needing to be determined.

IV. Validating Unit Commitment Models

Test data is important to not only independently validated the different UC models, but to compare them with one another accurately. The wide arrange of models and algorithms summarized were created to provide benefit to the TSOs,
whether it be reduce costs, decrease solution times, problem complexity, etc. The lack of system data being provided can make it challenging to validate a model’s real-world value. While most of the UC models and algorithms summarized in the previous section include some sort of test system and/or comparison to previous works, they lack complete system data (e.g., initial conditions or ramp rates), and validate the results with a single load profile. Initial conditions, end of simulation decisions, and “hardness” of the demand profile can create drastic changes in the optimal solution, result feasibility, and solution times. This doesn’t lead to a conclusive validation, especially for radically altered models or constraint implementations. It is unclear which constraints have been applied (validated) and difficult to recreate the results for future comparisons. A standard benchmark test system could help in validating UC models and provide a standard system to compare the performance between different models, algorithms and constraint configurations.

A UC benchmark test system should include a variety of generators and storage, varying network constraints, and multiple demand profiles (with a sub-hourly time-step) varying in “hardness” and allowing for a rolling UC to test and ensure all constraints and their implementations are validated. Complete generator data should be included, e.g., initial conditions, min. and max. generation limits, min. and max. up and down times, production costs, detailed start-up and shut-down costs, etc. This variation in units and load profiles is similar to the variety provided in the distribution test feeders [17]–[19] allowing for all types and configurations of components to be tested.

Fig. 1 plots a few different load profiles, that can be part of a benchmark test system. Note, a sub-hourly time-step is given to allow for future testing of sub-hourly models. Using a single load profile for comparing and validating a UC model or algorithm is not sufficient. The load profiles should be a combination of demand profiles and created test signals to challenge the UC models and ensure the constraints and their implementation are working as intended. For example, the sharp and smooth profiles both test the ramping capability and constraints, but with different levels of jumps in demand between time-steps. Running the same UC model with different load profiles can result in run times varying from seconds or less to hours. Even optimizing over sequential days, the “hardness” of the profile can alter significantly, thereby increasing or decreasing the difficulty of obtaining the optimal solution. Modifying any of the system parameters, e.g., initial conditions, ramp rates, etc., can compound these differences. Larger more complex MILP models are optimized until either the optimality gap or the maximum simulation time has been reached. Cases with easy load profiles can give a false since of a models performance (not the general optimal solution). As stated previous small changes in the solution can have significant effects in its operational implementation and cost in millions of dollars annually.

V. Conclusion

This paper provides a quick overview of a basic UC model along with a summary of the most cited works to help illustrate the need for benchmark test systems similar to those in other areas of power systems and operations research [15]–[21]. All the works presented offer some improvement to the UC problem, but their comparison and validation for real world use is difficult due to the lack of data that is typically not provided. Several examples are given, demonstrating how just changing basic parameters in the model or different load profiles can greatly effect the model’s feasibility, solution and solution time. As noted specifically for the PJM and CAISO systems, small
percentage changes in the solution can save or cost the TSO hundreds of millions of dollars per year [1].

ACKNOWLEDGMENT

This work was conducted in the Electricity Research Centre, University College Dublin, Ireland, which is supported by the Electricity Research Centres Industry Affiliates Programme (http://erc.ucd.ie/industry/). This material is based upon works supported by the Science Foundation Ireland, by funding Alexander C. Melhorn and Minsong Li, under Grant No. SFI/09/SRC/E1780. The opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Science Foundation Ireland.

REFERENCES


