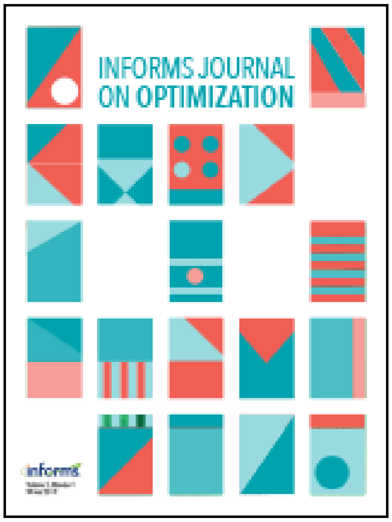


INFORMS Journal on Optimization

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Computationally Efficient Benders Decomposition for Energy Systems Planning Problems with Detailed Operations and Time-Coupling Constraints

Anna Jacobson, Filippo Pecci, Nestor Sepulveda, Qingyu Xu, Jesse Jenkins

INFORMS JOURNAL ON OPTIMIZATION
Articles in Advance, pp. 1–14
ISSN 2575-1484 (print), ISSN 2575-1492 (online)
https://pubsonline.informs.org/journal/ijoo

A Computationally Efficient Benders Decomposition for Energy Systems Planning Problems with Detailed Operations and Time-Coupling Constraints

Anna Jacobson,[a,*] Filippo Pecci,[b] Nestor Sepulveda,[c,d] Qingyu Xu,[e] Jesse Jenkins[f]

[a] Lewis-Sigler Institute of Integrative Genomics, Princeton University, Princeton, New Jersey 08540; [b] Andlinger Center for Energy and the Environment, Princeton University, Princeton, New Jersey 08540; [c] Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; [d] McKinsey and Company, New York, New York 10007; [e] Energy Internet Research Institute, Tsinghua University, Beijing 100084, People's Republic of China; [f] Andlinger Center for Energy and the Environment and Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08540

*Corresponding author

Contact: anna.f.jacobson@gmail.com, https://orcid.org/0000-0002-3929-4742 (AJ); fp0820@princeton.edu, https://orcid.org/0000-0003-3200-0892 (FP); nsep@mit.edu, https://orcid.org/0000-0003-2735-8769 (NS); xuqingyu0610@gmail.com, https://orcid.org/0000-0003-2692-5135 (QX); jdj2@princeton.edu, https://orcid.org/0000-0002-9670-7793 (JJ)

Received: February 10, 2023
Revised: April 18, 2023
Accepted: June 16, 2023
Published Online in Articles in Advance: August 2, 2023

https://doi.org/10.1287/ijoo.2023.0005

Abstract. Energy systems planning models identify least-cost strategies for expansion and operation of energy systems and provide decision support for investment, planning, regulation, and policy. Most are formulated as linear programming (LP) or mixed integer linear programming (MILP) problems. Despite the relative efficiency and maturity of LP and MILP solvers, large scale problems are often intractable without abstractions that impact quality of results and generalizability of findings. We consider a macro-energy systems planning problem with detailed operations and policy constraints and formulate a computationally efficient Benders decomposition separating investments from operations and decoupling operational timesteps using budgeting variables in the master model. This novel approach enables parallelization of operational subproblems and permits modeling of relevant constraints coupling decisions across time periods (e.g., policy constraints) within a decomposed framework. Runtime scales linearly with temporal resolution; tests demonstrate substantial runtime improvement for all MILP formulations and for some LP formulations depending on problem size relative to analogous monolithic models solved with state-of-the-art commercial solvers. Our algorithm is applicable to planning problems in other domains (e.g., water, transportation networks, production processes) and can solve large-scale problems otherwise intractable. We show that the increased resolution enabled by this algorithm mitigates structural uncertainty, improving recommendation accuracy.

Funding: Funding for this work was provided by the Princeton Carbon Mitigation Initiative (funded by a gift from BP) and the Princeton Zero-carbon Technology Consortium (funded by gifts from GE, Google, ClearPath, and Breakthrough Energy).
Supplemental Material: The e-companion is available at https://doi.org/10.1287/ijoo.2023.0005.

Keywords: macro-energy systems • capacity expansion planning • Benders decomposition • mixed integer linear programming • linear programming • decomposition methods

1. Introduction

Energy systems planning problems optimize resource investments, retirements, and operations to minimize total system cost (equivalently, maximize societal welfare) subject to technological, political, environmental, and economic constraints. These capacity expansion problems support decision-making in investment planning, regulation, and policy. Assuming perfect foresight and free entry, central-planning optimizations simulate a market under perfect competition; they are thus able to go beyond providing guidance on capacity deployment and retirement and play essential roles analyzing government policies (Larson et al. 2021, Victoria et al. 2022, Ricks et al. 2023) and advanced technologies' role in decarbonized energy systems (Mallapragada et al. 2020, Victoria et al. 2020, Ricks et al. 2022).

The vast majority of macro-energy systems planning problems (large-scale planning problems with regional or national scope) are implemented as linear programming (LP) or mixed integer linear programming (MILP) problems (Ringkjøb et al. 2018, Cho et al. 2022) due to the relative simplicity, computational efficiency, and maturity of LP and MILP solution methods and the fact that most salient system characteristics can be represented with reasonable accuracy using linear formulations. Recently, the increasing penetration of variable renewable energy resources (VRE) has required much greater temporal, geospatial, and operational resolution to accurately capture key physical, economic, and engineering considerations that affect investment and retirement decisions (Helistö et al. 2021). Electrification of transportation, heating, and industrial processes and production of hydrogen from electrolysis is also increasing the relevance of electricity systems and more tightly-coupling electricity with other energy and industrial systems, requiring the formulation and solution of multisector energy systems models (Brown et al. 2018, He et al. 2021). As a result, full-resolution, full-scale models are composed of millions of variables and constraints, thereby risking intractability.

In order to deal with computational constraints, macro-energy systems planning problems are often heavily abstracted; most models either by downsample or subsample constituent time periods, aggregate regions into larger geographic zones, and/or relax operational constraints of physical systems. For a recent review, see Cho et al. (2022).

Simplified model structures can improve runtime but significantly impact investment and policy recommendations derived from energy planning models (Bistline et al. 2022). Modeling too few representative days may fail to capture weekly demand patterns and their influence on thermal plant unit commitment (UC) and storage dispatch decisions (Mallapragada et al. 2020). Employing nonsequential time slices may poorly incorporate weather patterns (Poncelet et al. 2016) and prevent accurate modeling of flexibility requirements or energy storage. Temporal clustering is shown to significantly impact investment and operation of VRE (Pfenninger 2017). Simplification of system operation has been shown to affect investment recommendations in transmission planning (Xu and Hobbs 2019, Neumann et al. 2022) and dispatch decisions for systems with UC constraints (Palmintier and Webster 2013, Poncelet et al. 2020). Different means of geospatial aggregation techniques also impact model output (Siala and Mahfouz 2019, Frysztacki et al. 2022). Because of these effects, abstractions must be carefully tailored to each study's focus; this inhibits the generalizability of any given model and the replicability of solutions when multiple models' solutions are compared.

Energy systems planning problems are often intractable even while deploying significant abstractions; these models require computationally efficient solution methods to terminate. Lohmann and Rebennack (2017) developed tailored Benders decomposition algorithms for three simplified planning problems. Among the cases studied, two did not include time coupling constraints, as this omission allowed decomposed subproblems to be solved in parallel once investment decision variables had been fixed. The third case included just a single aggregated demand zone and omitted transmission operations and investments, storage resources, and policy constraints. Multiperiod planning problems solved in Lara et al. (2018) and Li et al. (2022) considered detailed operational and time-coupling constraints and decomposed the resulting MILP problems into a series of optimization problems per planning period. These studies did not investigate computationally efficient methods to solve the single-period subproblems; as a result, the largest case study in Lara et al. (2018) and Li et al. (2022) had only 6 zones and modeled each planning year using only 15 representative days. The work by Munoz et al. (2016), is one of the few to investigate computationally efficient methods for single-period planning problems. This study considered a full operational year with hourly resolution but ignored storage resources, ramping limits, and UC constraints, all of which couple operational decisions across time periods. Sepulveda (2020) proposed a nested decomposition algorithm to solve a single-year planning problem with detailed operational and time coupling constraints. In the first decomposition stage, investment decisions are separated from operational decisions. Then, the operational subproblem is solved using a Dantzig-Wolfe decomposition. However, such a technique does not allow for cuts to be fully decoupled by timestep and requires each iteration of the outer decomposition algorithm to await convergence of the inner decomposition before continuing to the next iteration.

Our study goes further than state-of-the-art models in the literature by investigating decomposition methods for a single-period energy systems planning problem with hourly resolution and detailed operation and time-coupling constraints. We aim to mitigate the errors introduced by temporal clustering by solving the planning problem for a full operational year with minimal downscaling or subsampling of intra-annual timesteps and to minimize abstractions of operational constraints. The model described below is an electricity system planning problem with detailed operational decisions and constraints on ramping, storage operations, and start-up and shut-down (UC) for thermal resources. Our formulation further includes policy constraints that couple time steps across the planning period, for example, caps on annual CO_2 emissions or a renewable portfolio standard (RPS). Decision variables consist of generation, energy storage and transmission investment and retirement decisions and operational decisions like resource dispatch. For MILP cases, we constrain all investment decisions to be integer. In comparison with previous literature (Lara et al. 2018, Li et al. 2022), we solve the resulting energy systems planning problem for a full operational year with hourly resolution, resulting in 8,736 time steps (52 weeks).

In this work, we propose a new decomposition scheme to separate investment decisions into a master model, along with a series of budgeting variables representing the time-coupling constraints (namely, CO_2 and RPS constraints.) This representation of policy constraints using budgeting variables is novel, to the best of the authors' knowledge. The reformulation described below enables the decomposition of a monolithic problem into a master model and several operational subproblems. Because each subproblem is fully decoupled, operational subproblems are solved in parallel with solutions incorporated as one Benders cut per subproblem per algorithmic iteration. As shown in Section 4.2, inclusion of multiple cuts per iteration significantly improves the method's computational performance compared with both monolithic solution approaches and standard, single-cut, Benders decomposition implementations. The inclusion of the budgeting variables for time-coupling constraints is thus a key novel contribution of this study.

The capacity expansion problems considered here belong to the wider framework of integrated planning in infrastructure systems. These optimization problems are characterized by a complex structure wherein diagonal blocks of the matrix composing systems' variables and constraints are linked by both complicating variables (e.g., investments) and complicating constraints (e.g., policy constraints) see Figure 1. Examples of other problems in this class include but are not limited to: optimal production planning in industrial processes, where inventory constraints link together different planning periods (Shah and Ierapetritou 2012); stochastic planning of water and wastewater systems (Naderi and Pishvaee 2017); and optimal placement of electric bus charging stations, where both locations and scheduling are optimized (An 2020). In all of these cases, the resulting optimization problems are too large to be solved by monolithic approaches and tailored decomposition methods such as the one described here are required. Because our decomposition scheme is not restricted to the framework of energy systems, our computational experiments can inform researchers working on integrated planning problems in other application areas.

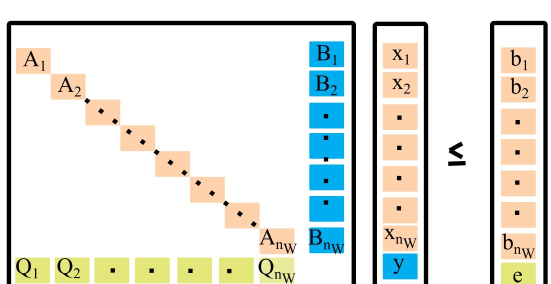
Section 2 describes the formulation of the energy systems planning problem with detailed operational and time-coupling constraints. In Section 3, we develop a novel Benders decomposition scheme which divides a full operational year into subperiods which can be processed in parallel. In Section 4, we evaluate the efficiency of the developed method using case studies derived from the Eastern United States with varying spatial extent ranging from 2–19 zones and with levels of temporal resolution ranging from 2 to 52 weeks. We conclude by examining the impacts of our enabled increased temporal resolution on resources' investment recommendations.

2. Problem Formulation

We formulate an energy systems planning problem wherein electricity generation, energy storage, capacity expansion and retirement, and energy dispatch are jointly optimized over a single planning period. An in-depth description of the constraints comprising the problem are included in the appendix, sections A2.1–A2.3. The objective function being minimized represents investment and operational costs and includes penalties for violating policy constraints. In this work, we consider three different policy scenarios and their corresponding policy constraints, as described in Table 1.

Optimization constraints (see section A2.2) include maximum output limits for renewable resources, ramping limits, and start-up and shut-down restrictions for thermal generators (e.g., UC constraints.) As noted by

Figure 1. (Color online) Block Structure of Problem (1) with Both Complicating Constraints and Variables, Where $n_W = |W|$



Note. Investment-only constraints (1d) are not pictured.

Table 1. Policy Scenarios Considered in this Study

Scenario	Description
REF	The reference case. Emissions and dispatch by resource type are unrestricted.
RPS	Renewable portfolio standard. 70% of generation must come from qualifying resources (e.g., VRE).
CO_2	CO_2 emissions cap. Emissions are constrained to 0.05 tons per MWh.

Palmintier and Webster (2015), the ability to represent UC significantly impacts model results and improves accuracy. Here, we implement an aggregated UC model similar to Palmintier and Webster (2013, 2015) and relax the integer constraints on UC variables to further improve scalability.

In most cases, planning problems for energy networks with several thousand nodes (i.e., buses) are intractable; we therefore divide the geographical area of interest into zones which incorporate real-world data (see section A2.5) on demand profiles and climate conditions. For each zone, we create clusters for resources (e.g., generators or storage units) based on technology, cost of connection to transmission grids, and operating parameters. We assume that resources within a given cluster have the same capacity size per unit and use integer variables to model the decisions to install or retire a number of units within each cluster. We also include interregional power transmission and integer investment decision variables for capacity expansion of existing transmission paths between modeled zones. In this way, the energy system is modeled as a graph where each node represents a zone with demand and constituent resources (e.g., generators, storage) and edges represent interzonal transmission capacity.

In this study, we implement a transport model and do not consider Kirchhoff's Voltage Law (KVL) or power losses. As noted in Neumann et al. (2022), inclusion of these features would have greater impact on systems with lower levels of spatial aggregation (e.g., hundreds of nodes) than those described in Section 4, which include up to 19 aggregated demand zones. However, the novel Benders decomposition method developed in Section 3 can be applied to planning problems with KVL constraints due to the separation of transmission investments and operations without the need for nonlinear solution methods; this is the subject of future work.

Macro-energy systems planning problems are commonly formulated using a selection of operational subperiods due to models' size (Frew and Jacobson 2016, Lara et al. 2018, Mallapragada et al. 2018, Mallapragada et al. 2020, Li et al. 2022). Storage resources' operation and thermal resource UC and ramping constraints are often modeled within each subperiod by linking first and last time steps in a method known as "circular indexing." This approximation assumes that storage levels and UC decisions across two subperiods are decoupled. Errors may arise when subperiods are too short (Mallapragada et al. 2020), as it becomes impossible to fully capture weekly demand and weather patterns and their influence on UC and storage dispatch decisions. This is often the case in previous studies (Lara et al. 2018, Mallapragada et al. 2018, Li et al. 2022), where only a few representative 24-hour periods (days) are used to model systems' yearly operation. For the work in this manuscript, subperiods are one week long; circular indexing occurs over a 168-hour time period.

We develop a Benders decomposition scheme to solve an optimization problem for a planning period with up to 52 consecutive weeks of operational decisions. The investigation of appropriate timeseries clustering methods to select representative subperiods is outside the scope of this manuscript. For problems with fewer than 52 weeks we assume that a clustering method has been implemented, resulting in sampled subperiods indexed by set W. For each subperiod $w \in W$, we define its set of hours as $H_w = \{(w-1)\delta_w + 1, \dots, w\delta_w\}$, where δ_w is the number of time steps within the subperiod (in our case, $\delta_w = 168$).

2.1. Overall Problem Formulation

Appendix Section A.2 contains a list of all constraints included in the model. To highlight problem structure, we introduce a compact formulation below. Assume that when $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors, the inequality $\boldsymbol{u} \le \boldsymbol{v}$ is evaluated component-wise. Let $\boldsymbol{y} \in \mathbb{R}^m$ be a vector grouping all investment decision variables, and let $\boldsymbol{R}$ and $\boldsymbol{r}$ be such that constraints (A1)–(A3) correspond to $\boldsymbol{R}\boldsymbol{y} \le \boldsymbol{r}$. In addition, let vector $\boldsymbol{c}_I$ be such that the fixed cost objective terms (A18) are denoted $\boldsymbol{c}_I^T \boldsymbol{y}$. For each subperiod $w \in W$, consider a vector $\boldsymbol{x}_w \in \mathbb{R}^n$ grouping all operational decision variables, and let matrices $\boldsymbol{A}_w$ and $\boldsymbol{B}_w$, and vector $\boldsymbol{b}_w$ be such that $\boldsymbol{A}_w \boldsymbol{x}_w + \boldsymbol{B}_w \boldsymbol{y} \le \boldsymbol{b}_w$ corresponds to constraints (A4)–(A13). Let vector $\boldsymbol{c}_w$ be such that the objective function terms (A19) + (A21) + (A20) + (A22) are equal to $\sum_{w \in W} \boldsymbol{c}_w^T \boldsymbol{x}_w$. Finally, let matrix $\boldsymbol{Q}_w$ and vector $\boldsymbol{e}$ be such that $\sum_{w \in W} \boldsymbol{Q}_w \boldsymbol{x}_w \le \boldsymbol{e}$ represents the policy constraints across the different scenarios (corresponding to (A15) for case RPS, to (A16) for case CO_2, and remaining unenforced for case REF.) The resulting MILP problem is:

$$\text{minimize} \quad \boldsymbol{c}_I^T \boldsymbol{y} + \sum_{w \in W} \boldsymbol{c}_w^T \boldsymbol{x}_w \tag{1a}$$

$$\text{subject to} \quad \boldsymbol{A}_w \boldsymbol{x}_w + \boldsymbol{B}_w \boldsymbol{y} \leq \boldsymbol{b}_w, \quad \forall w \in W \tag{1b}$$

$$\sum_{w \in W} \boldsymbol{Q}_w \boldsymbol{x}_w \leq \boldsymbol{e} \tag{1c}$$

$$\boldsymbol{R}\boldsymbol{y} \leq \boldsymbol{r} \tag{1d}$$

$$\boldsymbol{x}_w \geq 0, \quad \forall w \in W \tag{1e}$$

$$\boldsymbol{y} \geq 0 \tag{1f}$$

$$\boldsymbol{y} \in \mathbb{Z}^m. \tag{1g}$$

3. Solution Method

Problem (1) is difficult to solve because it includes both complicating variables (e.g., investment decisions) and constraints (e.g., CO_2 limits) that link all operational subperiods; see Figure 1.

The complex structure of Problem (1) is shared by several integrated planning problems in a variety of application areas (e.g., Shah and Ierapetritou 2012, Naderi and Pishvaee 2017, Lara et al. 2018, An 2020). While ubiquitous, analogous optimization problems are particularly common in macro-energy systems modeling. Multiperiod planning problems in Lara et al. (2018) and Li et al. (2022) are decomposed into a series of single-period operational problems that are a special case of (1), where investment decision variables are fixed. In these studies, the single-period operational problems were not further decomposed due to the presence of policy constraints like those seen in (A15) and (A16). Few studies have developed efficient methods for single-period planning problems: Munoz et al. (2016) investigated the solution of a special case of Problem (1) where storage, ramping limits, and thermal plant UC are not considered. However, the exclusion of UC as a system characteristic affects the accuracy of model results (Palmintier and Webster 2015). Studies like the one performed by Lohmann and Rebennack (2017) included simplified cases of Problem (1), where either time-coupling or interzonal transmission constraints are ignored.

In the following, we develop a Benders decomposition scheme for Problem (1) which enables parallel computation of subperiods. A standard implementation of Benders decomposition to Problem (1) would consider only investment decisions $\boldsymbol{y}$ as complicating variables, and rely on the solution of a large, monolithic operational subproblem, spanning the whole year, at each iteration. This is the same approach used in Lara et al. (2018) and Li et al. (2022) to solve analogous energy systems planning problems. Nested decompositions have also been proposed (Sepulveda 2020) as means of further decomposing operational subproblems, but this technique still takes the operational problem as a single entity in the scope of the master model. In both cases, the master problem is given by:

$$\text{minimize} \quad \boldsymbol{c}_I^T \boldsymbol{y} + \theta \tag{2a}$$

$$\text{subject to} \quad \theta \geq \sum_{w \in W} f_w^j + (\boldsymbol{\pi}^j)^T (\boldsymbol{y} - \boldsymbol{y}^j), \quad \forall j = 0, \ldots, k-1, \tag{2b}$$

$$\boldsymbol{R}\boldsymbol{y} \leq \boldsymbol{r} \tag{2c}$$

$$\boldsymbol{y} \geq 0 \tag{2d}$$

$$\boldsymbol{y} \in \mathbb{Z}^m. \tag{2e}$$

Where θ represents an approximation of the operational cost within the master model. Benders cuts (2b) approximate operational cost based on a single operational subproblem, which models dispatch across the entire time-series:

$$\text{minimize} \quad \sum_{w \in W} \boldsymbol{c}_w^T \boldsymbol{x}_w \tag{3a}$$

$$\text{subject to} \quad \boldsymbol{A}_w \boldsymbol{x}_w + \boldsymbol{B}_w \boldsymbol{y} \leq \boldsymbol{b}_w, \quad \forall w \in W \tag{3b}$$

$$\sum_{w \in W} \boldsymbol{Q}_w \boldsymbol{x}_w \leq \mathbf{e} \tag{3c}$$

$$\boldsymbol{x}_w \geq 0, \quad \forall w \in W \tag{3d}$$

$$\boldsymbol{y} = \boldsymbol{y}^k \qquad : \boldsymbol{\pi} \tag{3e}$$

Problem (3) is not separable with respect to the subperiod index $w \in W$ due to constraints in (3c) that tie together all subperiods. As shown in Section 4, modeling operations as a monolithic problem is impractical as Problem (3) quickly becomes intractable as number of zones and subperiods increases. Sepulveda (2020) suggested solving Problem (3) using a nested Dantzig-Wolfe decomposition algorithm. However, such a strategy requires each iteration of the Benders decomposition algorithm to await convergence of the inner Dantzig-Wolfe decomposition before continuing. It also does not allow the inclusion of multiple decoupled cuts (2b) per iteration, resulting in increased number of iterations as seen in Table 4.

In the following we take a different approach, enabling the separation of Problem (1) into subproblems that can be solved in parallel without requiring a nested decomposition. First, we prove the following result.

Theorem 1. *Problem* (1) *is equivalent to:*

$$\begin{aligned}
\textit{minimize} \quad & \boldsymbol{c}_I^T \boldsymbol{y} + \sum_{w \in W} \boldsymbol{c}_w^T \boldsymbol{x}_w \\
\textit{subject to} \quad & \boldsymbol{A}_w \boldsymbol{x}_w + \boldsymbol{B}_w \boldsymbol{y} \le \boldsymbol{b}_w, \quad \forall w \in W \\
& \boldsymbol{Q}_w \boldsymbol{x}_w \le \boldsymbol{q}_w, \quad w \in W \\
& \sum_{w \in W} \boldsymbol{q}_w = \boldsymbol{e} \\
& \boldsymbol{R}\boldsymbol{y} \le \boldsymbol{r} \\
& \boldsymbol{x}_w \ge 0, \quad \forall w \in W \\
& \boldsymbol{y} \ge 0 \\
& \boldsymbol{y} \in \mathbb{Z}^m.
\end{aligned} \tag{4}$$

Proof. Without loss of generality, we assume that $W = \{1, \dots, n_W\}$. We will show that:

$$\sum_{i=1}^{n_W} \boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{e} \Longleftrightarrow \exists (\boldsymbol{q}_i)_{i=1}^{n_W} \text{ such that } \sum_{i=1}^{n_W} \boldsymbol{q}_i = \boldsymbol{e} \text{ and } \boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{q}_i, \quad \forall i = 1, \dots, n_W. \tag{5}$$

The implication "$\Longleftarrow$" follows by definition of vectors $\boldsymbol{q}_1, \dots, \boldsymbol{q}_{n_W}$. By induction, we show that also the "$\Rightarrow$" implication holds. Assume $n_W = 2$ and $\boldsymbol{Q}_1\boldsymbol{x}_1 + \boldsymbol{Q}_2\boldsymbol{x}_2 \le \boldsymbol{e}$. Then, if we define $\boldsymbol{q}_1 = \boldsymbol{e} - \boldsymbol{Q}_2\boldsymbol{x}_2$ and $\boldsymbol{q}_2 = \boldsymbol{Q}_2\boldsymbol{x}_2$ we have

$$\begin{aligned}
& \boldsymbol{q}_1 + \boldsymbol{q}_2 = \boldsymbol{e} \\
& \boldsymbol{Q}_1\boldsymbol{x}_1 \le \boldsymbol{e} - \boldsymbol{Q}_2\boldsymbol{x}_2 = \boldsymbol{q}_1 \\
& \boldsymbol{Q}_2\boldsymbol{x}_2 = \boldsymbol{q}_2
\end{aligned} \tag{6}$$

Next, we assume that it holds for $n_W = l$ and prove that this implies that it holds for $n_W = l + 1$. We want to show that:

$$\sum_{i=1}^{l+1} \boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{e} \Rightarrow \exists \boldsymbol{q}_1, \dots, \boldsymbol{q}_{l+1} \text{ such that } \sum_{i=1}^{l+1} \boldsymbol{q}_i = \boldsymbol{e} \quad \text{and} \quad \boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{q}_i, \ \ \forall i = 1 \dots, l+1 \tag{7}$$

Define $\boldsymbol{q}_{l+1} = \boldsymbol{Q}_{l+1}\boldsymbol{x}_{l+1}$ and observe:

$$\sum_{i=1}^{l} \boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{e} - \boldsymbol{Q}_{l+1}\boldsymbol{x}_{l+1} = \boldsymbol{e} - \boldsymbol{q}_{l+1}. \tag{8}$$

Since we are assuming that the claim holds for $n_W = l$, we have that there exist $\boldsymbol{q}_1, \dots, \boldsymbol{q}_l$ such that $\sum_{i=1}^{l} \boldsymbol{q}_i = \boldsymbol{e} - \boldsymbol{q}_{l+1}$ and $\boldsymbol{Q}_i \boldsymbol{x}_i \le \boldsymbol{q}_i$ for all $i = 1, \dots, l$, which completes the induction step. □

3.1. Benders Decomposition Algorithm

The structure of Problem (4) suggests the implementation of a Benders decomposition algorithm wherein budgeting variables ($\boldsymbol{q}_w$) are used to implement (1c) and both $\boldsymbol{y}$ and $\boldsymbol{q}_w$ are treated as complicating variables. We note that this budgeting approach should be extensible to other constraints coupling subperiods, such as long-duration storage state-of-charge or hydro reservoir levels or similar time-coupling inventory constraints, although these constraints are not implemented in the present numerical tests. Initialize $\text{UB} = \infty, f_w^0 = 0$, $\boldsymbol{\pi}^0 = \boldsymbol{0}$, $\boldsymbol{y}^0 = \boldsymbol{0}$, $\boldsymbol{\lambda}_w^0 = \boldsymbol{0}$, and $\boldsymbol{q}_w^0 = \boldsymbol{0}$, for all $w \in W$. For all $k = 1, \dots K_{\max}$ proceed as follows:

Step 1. Obtain estimated solutions $\boldsymbol{y}^k$ and $\boldsymbol{q}_w^k$, for all $w \in W$ by solving the master problem:

$$\text{minimize} \quad \boldsymbol{c}_I^T \boldsymbol{y} + \sum_{w \in W} \theta_w \tag{9a}$$

$$\text{subject to} \quad \theta_w \geq f_w^j + (\boldsymbol{\pi}^j)^T(\boldsymbol{y} - \boldsymbol{y}^j) + (\boldsymbol{\lambda}_w^j)^T(\boldsymbol{q}_w - \boldsymbol{q}_w^j), \quad \forall j = 0, \ldots, k-1, \, w \in W \tag{9b}$$

$$\sum_{w \in W} \boldsymbol{q}_w = \boldsymbol{e} \tag{9c}$$

$$\boldsymbol{R}\boldsymbol{y} \leq \boldsymbol{r} \tag{9d}$$

$$\boldsymbol{y} \geq 0 \tag{9e}$$

$$\boldsymbol{y} \in \mathbb{Z}^m, \tag{9f}$$

where f_w^j represents the current operational cost given a fixed set of investments and $\boldsymbol{\pi}^j$ and $\boldsymbol{\lambda}_w^j$ represent the Lagrangian multipliers associated with investment decisions and policy constraints (e.g., RPS) respectively. Set LB to be the optimal value of Problem (9).

Step 2. For every $w \in W$, solve the following linear subproblem:

$$\begin{array}{lll} \text{minimize} & \boldsymbol{c}_w^T \boldsymbol{x}_w & \\ \text{subject to} & \boldsymbol{A}_w \boldsymbol{x}_w + \boldsymbol{B}_w \boldsymbol{y} \leq \boldsymbol{b}_w & \\ & \boldsymbol{Q}_w \boldsymbol{x}_w \leq \boldsymbol{q}_w & \\ & \boldsymbol{x}_w \geq 0 & \\ & \boldsymbol{y} = \boldsymbol{y}^k & : \boldsymbol{\pi} \\ & \boldsymbol{q}_w = \boldsymbol{q}_w^k & : \boldsymbol{\lambda}, \end{array} \tag{10}$$

and compute the optimal value f_w^k and Lagrangian multipliers $\boldsymbol{\pi}^k$ and $\boldsymbol{\lambda}_w^k$ given the fixed set of investment decisions ($\boldsymbol{y}^k$) and coupling constraint budgets ($\boldsymbol{q}_w^k$) which are constants in the scope of the subproblem. Set $\text{UB} = \min(\text{UB}_{\text{prev}}, \boldsymbol{c}_I^T \boldsymbol{y}^k + \sum_{w \in W} f_w^k)$, where UB_{prev} is the upper bound from the previous iteration of the algorithm, (∞ if $k = 1$.) If $(\text{UB} - \text{LB})/\text{LB} \leq \texttt{Rel}_{tol}$, then stop. Else, set $k = k + 1$ and go back to Step 1.

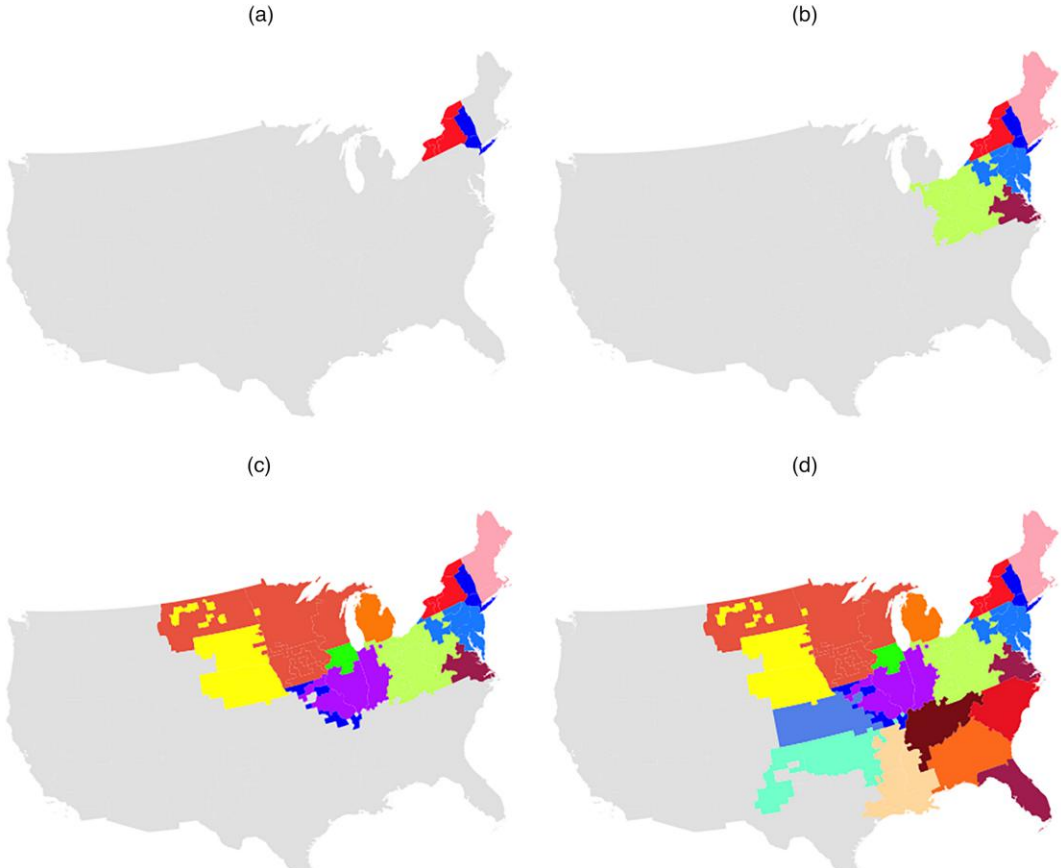
Observe that subproblems in Step 2 are separable and can be solved in parallel, solving $|W|$ smaller operational subproblems, each formulated over δ_w timesteps. This comes at the cost of adding $|W|$ constraints (Benders cuts) to the master problem per iteration, thereby increasing the size of the master problem compared with a standard Benders implementation. Numerical experiments reported in the next section suggest that this trade-off is worthwhile, as the total computational effort is dominated by the solution of the operational subproblems rather than the master problem. Furthermore, incorporation of multiple cuts allows more information to be communicated to the master problem at each iteration of the algorithm, decreasing the number of iterations needed for convergence.

4. Numerical Experiments

We consider 2-, 6-, 12-, and 19-zone cases of the Eastern United States; see Figure 2. Initial capacity estimates are given for the year 2050 as output by the case generation software PowerGenome (Schivley et al. 2021), see section A2.5. For each set of Eastern Interconnection regions, we represent the operational year using 2, 12, 22, 32, 42, and 52 representative weeks. Each planning problem is considered under the three policy scenarios described in Table 1. The size of the resulting optimization problems is summarized in Table 2.

In order to evaluate the impact of integer investment decision variables on the computational effort needed to solve energy systems planning problems, we consider two forms of Problem (1): in the first (hereafter "MILP") investment and retirement variables are integer for both generation and transmission resources—this is the form presented in Problem (1). In the second (hereafter "LP") we relax this constraint and allow all variables to be continuous. We compare the computational performance of the decomposition algorithm with a monolithic approach, where Problem (1) is solved directly by state-of-the-art optimization solvers, and with a standard Benders decomposition algorithm where a single full operational subproblem is solved at each iteration.

Figure 2. (Color online) IPM Regions Used in Our Numerical Experiments



Notes. (a) Zones, 2-zone simulation; (b) Zones, 6-zone simulation. (c) Zones, 12-zone simulation; (d) Zones, 19-zone simulation.

4.1. Computational Setup

All optimization problems are implemented in Julia 1.6.1 (Bezanson et al. 2017), where optimization solvers are called through JuMP 0.21.9 (Dunning et al. 2017). We solve all MILP and LP problems using Gurobi (v9.0.1). The simulations are run on Princeton University's Della computer cluster with 2.8 GHz Cascade Lake processors, Intel Xeon Platinum 8260 at 2.40 GHz. Problems are considered intractable if they require more than 48 hours of computations or 200 GB of memory to terminate.

We set a tolerance $\mathtt{Rel}_{tol} = 10^{-3}$. Accordingly, optimality tolerance, MIP gaps, and barrier convergence tolerance for the monolithic models are set to 10^{-3}. For these models, Gurobi is run using the barrier method with crossover turned off. This allows a fair comparison between the runtimes of monolithic and decomposed approaches, as Benders decomposition is not guaranteed to return basic solutions, and the crossover stage of the interior point solution method implemented by Gurobi requires considerable computational time to convert an optimal feasible solution to an optimal basic solution. When solving the subproblem (10), however, we enable crossover, as cuts (9b) require basic solutions.

Table 2. Model Size by Zone

Zones	$\lvert G \rvert$	$\lvert G^{UC} \rvert$	Variables ($\cdot 10^9$)	Constraints ($\cdot 10^9$)
2	62	16	1.1	3.4
6	175	54	3.4	10.5
12	285	106	6.2	19.3
19	437	167	9.7	30.4

Notes. Shows number of generator clusters, total number of variables and total number of constraints. Numbers of variables and constraints are those associated with the 52-week (full year) monolithic problem.

We solve the operational subproblems (10) in parallel. For each iteration of the algorithm, $|W|$ CPUs are assigned one operational subproblem each. Data pertaining to cuts in (9b) is returned to the master model from these auxiliary CPUs. Returning only information necessary to formulate cuts (as opposed to entire solutions) allows minimal interprocess communication, thus speeding up time between iterations. Problem (9) waits for all subproblems to terminate before incorporating each of their cuts into a new iteration of the master problem and sending solutions $\boldsymbol{y}^k$ and $\boldsymbol{q}_w^k$ back to the subproblems for the next algorithmic iteration. For these cases, all subproblems were solved on a single compute node. In cases where the number of subproblems exceeds available CPUs on a single compute node, multiple subproblems may be run in series on each CPU, or multiple nodes could be coordinated via distributed parallelization (which incurs greater computational overhead for interprocess coordination across nodes).

4.2. Results

We define runtime as the time spent initializing the problem, performing all computations within solvers, and writing solutions to file at the end of simulations. Table 3 compares the computational performance of the Benders decomposition algorithm and the monolithic solution approach.

We find that using Benders decomposition brings little benefit for cases with small numbers of zones, especially for LP problems without integer variables. We observe a substantial reduction in computational time when considering MILP problems, particularly large scale problems. Additionally, all instances of the Bender's algorithm solved within the time limit of 48 hours whereas the monolithic algorithm saw 29 intractable MILP cases out of the 70 that were run.

We observe that most intractable problem instances for the monolithic model occur in scenarios RPS and CO_2. Including these time-coupling constraints reflects common clean energy and emissions reductions targets; in our trials, the reference, RPS-, and CO_2-constrained cases incurred 8.6e8, 6.0e8, and 1.9e8 tons of CO_2, respectively, across the year for the 19-zone case when simulated on 12 weeks of data. An inability to include these time-coupling constraints thus threatens models' relevance to systems undergoing decarbonization.

Figure 3 suggests that the runtime associated with our Benders algorithm grows linearly with the number of weeks considered. We observe that runtime grows quadratically with the number of modeled zones. This fact hints that there may be value in decomposing along spatial dimensions in future work, as the algorithm described here operates only over temporal dimensions.

As the number of weeks in the simulation increases, runtime per iteration increases linearly—see Figure 4. Because the master model incorporates more information per iteration due to its increased number of cuts, the number of iterations required for convergence decreases roughly logarithmically (Figure 4). These trends explain the superior performance of the decomposition method for problems with a larger numbers of weeks.

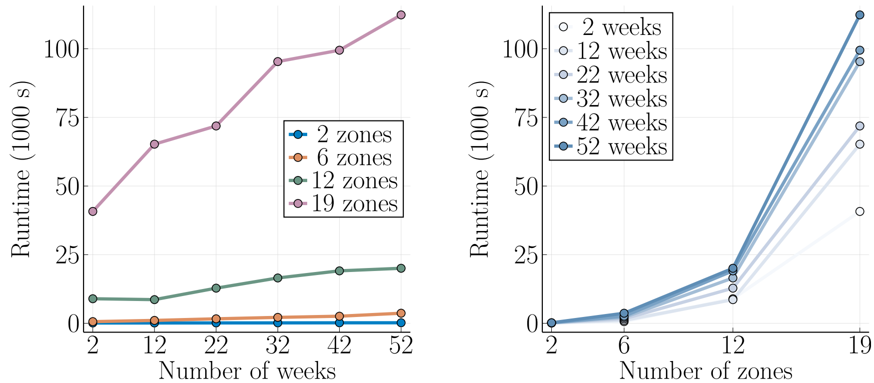
In contrast, increasing the number of zones increases the size of both the master (9) and the subproblem (10); accordingly, we see a simultaneous increase in iteration time and number of iterations required for convergence

Table 3. Runtime for Benders and Monolithic Models (100 s) Followed by Ratio Between Monolithic and Decomposition Solution Approaches

			LP						MILP					
Weeks →	$\|Z\|$	$\|G\|$	2	12	22	32	42	52	2	12	22	32	42	52
	2	62	1.1	1.2	**2.1**	**3.4**	**4.1**	**6.4**	**1.1**	**1.2**	**1.4**	**1.6**	**1.9**	**2.0**
	6	175	5.9	10.1	16.6	19.7	**26.0**	**28.3**	6.1	**10.6**	**16.4**	**21.6**	**25.9**	**36.8**
	12	285	66.9	79.3	112.8	135.1	151.6	173.0	89.5	**86.2**	**128.0**	**165.2**	**191.1**	**200.4**
Benders (100 s)	19	437	474.0	465.0	558.3	702.0	776.5	720.0	407.5	**652.4**	**718.9**	**953.3**	**994.7**	**1,123.4**
	2	62	**0.4**	**1.0**	2.1	3.5	4.1	6.0	2.9	58.6	112.5	702.8	1,344.3	651.9
	6	175	**0.6**	**4.3**	**10.8**	**18.3**	27.5	36.7	7.5	129.4	∞^T	∞^M	∞^T	∞^T
	12	285	**1.2**	**10.7**	**27.9**	**50.4**	**79.3**	**106.8**	**24.5**	∞^T	∞^T	∞^T	∞^T	∞^T
Mono. (100 s)	19	437	**1.7**	**21.7**	**61.9**	**112.0**	**167.9**	**227.4**	**252.5**	∞^T	∞^T	∞^T	∞^T	∞^T
	2	62	0.3	0.8	**1.5**	**2.2**	**2.3**	**3.2**	**2.5**	**48.0**	**78.2**	**444.8**	**722.7**	**329.2**
	6	175	0.1	0.4	0.7	0.9	**1.1**	**1.3**	**1.2**	**12.3**	∞^T	∞^M	∞^T	∞^T
	12	285	<0.1	0.1	0.3	0.4	0.5	0.6	0.3	∞^T	∞^T	∞^T	∞^T	∞^T
Ratio	19	437	<0.1	0.1	0.1	0.2	0.2	0.3	0.6	∞^T	∞^T	∞^T	∞^T	∞^T

Notes. ∞ denotes a case with an intractable monolithic model. Cases that are intractable due to memory are noted with the superscript M. Cases that are intractable due to insufficient time are noted with a superscript T. Cases where the model outperformed its analogous model formulation are shown for the runtime rows, cases where the decomposed model outperformed monolithic are bolded for the ratio cases. Results shown for the CO_2-constrained case. Additional cases are included in the appendix.

Figure 3. (Color online) Runtime by Weeks (Left) and by Zone (Right) Obtained by Applying Benders Decomposition to Solve the MILP Problem in the CO_2-Constrained Case



Note. Runtime grows linearly with the number of weeks, while it increases quadratically with the number of zones.

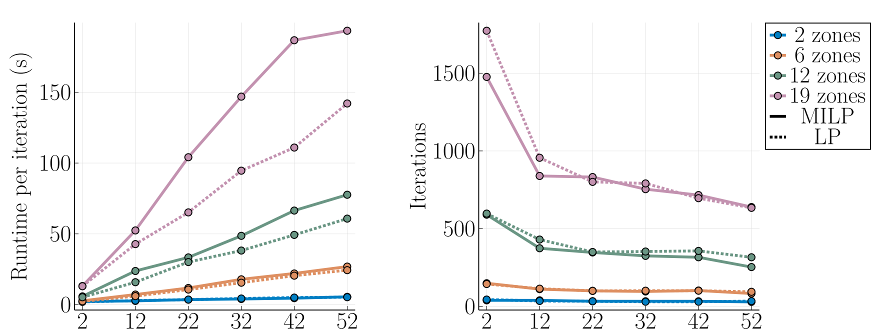
(Figure 4). These dual impacts explain the quadratic runtime increase shown in Figure 3 as the number of zones increases. Increasing the number of zones increases number of resource operational decisions and thus increases the amount of information that must be incorporated by a given cut, as the vector $\boldsymbol{y}$ in (9b) grows in dimensionality. The direct relationship between zones and number of iterations suggests that higher-dimensionality cuts poorly approximate the subproblems' objective values. Our decomposition method outperforms the monolithic approach on all MILP formulations with more than 2 weeks. It outperformed LP formulations on 2- and 6-zone problem instances with more than 12 or 32 weeks, respectively. The decomposition method remained within an order of magnitude of runtime on all simulations and always maintained a linear relationship between weeks and runtime, suggesting that the decomposed model will eventually outperform a monolithic approach as temporal resolution increases.

We further investigate the benefits of our novel decomposition scheme when compared with a standard implementation of Benders decomposition algorithm, which keeps timesteps coupled when solving Problem (1). In this case, the master problem is given by Equation (2), while the single operational subproblem is formulated as in (3).

The Benders algorithm with a fully coupled operational subproblem is intractable under all instances with more than 22 weeks of operation (Table 4). An aggregated Benders cut does not reduce the number of iterations as temporal resolution increases and further slows iteration time due to the size of the subproblem and the impossibility of parallelizing it.

Because it is the largest simulation found to be tractable under this standard Benders implementation, we consider a 6-zone LP network run on 22 weeks of data under the CO_2 policy scenario. Figure 5 compares the performance of the decomposition scheme developed in Section 3 with the standard Benders implementation as

Figure 4. (Color online) Runtime per Iteration by Week (Left) and Number of Iterations (Right), Shown for Different Numbers of Zones in the Reference Case



Notes. Runtime per iteration increases both with the number of weeks and the number of zones modeled. Number of iterations increases with the number of zones but decreases with the number of weeks modeled.

Table 4. Runtime in Seconds and Number of Iterations for 6-Zone LP CO_2-Constrained Simulations

Model Type	Runtime (100 s)						Iterations					
Weeks →	2	12	22	32	42	52	2	12	22	32	42	52
Monolithic	0.6	4.3	10.8	18.3	27.5	36.9	NA					
Benders	5.9	10.1	16.6	19.7	26.0	28.3	231	139	132	104	107	94
Benders full operation	40.6	5.7e2	1.5e3	INT	INT	INT	509	607	630	INT	INT	INT

Notes. 6-Zone CO_2 was selected for representation because it was the largest case with tractable full operation models. Intractable simulations are marked INT.

described in Equations (2) and (3). For both algorithms, we define the optimality gap as (UB − LB)/LB, where UB and LB are current upper and lower bounds on the optimal value, respectively. Figure 5 shows that the novel decomposition algorithm takes significantly less time to converge, running in 132 iterations with an average iteration time of 12.0 seconds. The standard Benders algorithm, by contrast, runs in 630 iterations with an average iteration time of 241.0 seconds (see Table 4).

The purpose of the energy systems planning models is to provide decision support for resource capacity expansion. As such, the specific solutions vector for these capacity planning problems is also of salient interest to stakeholders and decision makers (e.g., in utility integrated resource planning processes). Because system cost is not the primary output of these models, we do not believe it appropriate to rely on it as the sole indicator of the quality of models' results. In fact, Palmintier and Webster (2013) note that error in cost does not correlate with error in other system characteristics depending on the means of clustering used or capacity installed for different resources in the system. It is thus impossible to extrapolate error in capacity decisions based solely on error in cost. Therefore, we define the mean squared error (MSE) in investment decisions as:

$$\mathrm{MSE}(j) = \frac{1}{|G|}\sqrt{\sum_{g\in G}(y_g^{Pj} - y_g^{P52})^2}, \tag{11}$$

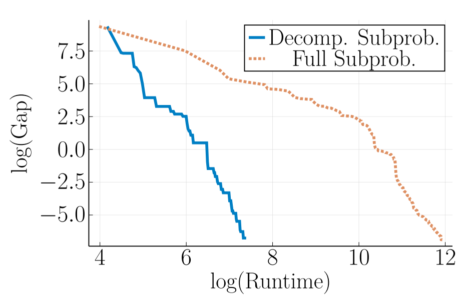
where y_g^{Pj} is the recommended capacity for resource $g \in G$, computed solving our planning problem for $j \in \{2, 12, 22, 32, 42, 52\}$ weeks.

These MSE values are computed by comparing near-optimal solutions for the considered MILPs. There may be multiple investment portfolios yielding similar costs for any given model. The MSE metric described in Equation (11) assesses the ability of lower-resolution models to recapitulate investment recommendations of higher-resolution models which inherently have lower structural uncertainty, even considering the possibility of multiple near-optimal solutions.

Figure 6 shows that MSE decreases as the temporal resolution increases and that using too few representative weeks can lead to average deviations as high as 600 MW per site for storage clusters. While we show that MSE decreases as temporal resolution increases, it is impossible to compute rigorous upper bounds on MSE a priori. That is, it is impossible to prove the maximum deviation from optimal capacity investments due to abstractions without solving the full model for comparison.

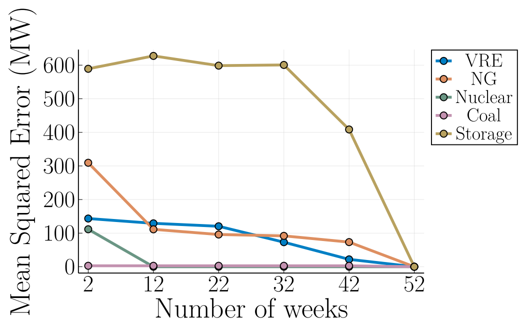
Figure 6 shows that temporal resolution has greatest impact on storage, NG, and VRE resources, likely due to the misrepresentation of VRE availability in models with low temporal resolution. The relatively high MSE

Figure 5. (Color online) Optimality Gap vs. Runtime for Our Benders Algorithm with Decomposed Subproblem (10) and a Standard Benders Implementation with a Full Operational Subproblem (3)



Note. Plotted on logarithmic x- and y-axes.

Figure 6. (Color online) Mean Squared Error as Defined in Equation (11)



Notes. Figure shows that low temporal resolution causes great deviation in capacity recommendations for storage, NG, and VRE. Data are shown for the 19-zone, CO_2-constrained MILP trial.

values for resources in Figure 6 demonstrate models' difficulty in providing recommendations on storage, NG, and VRE investments at high levels of temporal aggregation.

Coal resources are incentivized to retire completely by policy constraints. Similarly, nuclear resources are disincentivized from building due to a high fixed cost of investment. These impacts yield little variation in capacity regardless of temporal resolution, leading to low MSE in Figure 6. Errors at the resource level are likely to vary with policy or cost assumptions.

5. Conclusions

Our algorithm was able to solve MILP energy systems planning problems that were intractable when monolithic; we showed that this ability markedly decreased error associated with modeling. Furthermore, Figure 5 shows that the standard Benders algorithm without decoupling via budgeting constraints in the master model is significantly less successful than our fully decomposed scheme, often leading to intractability on smaller problems than the monolithic approach (Table 4). Standard decomposition schemes that do not separate timesteps do not gain as much information from Benders cuts at each iteration. We therefore note that the reformulation of the problem from (1) to (4) is the primary contribution of this paper.

The proposed decomposition scheme has several advantages over monolithic models currently in use industry-wide and methods explored in previous literature where operational subproblems were not decomposed (Lara et al. 2018, Li et al. 2022) or were decomposed via a nested algorithm (Sepulveda 2020). Below, we list some specific benefits to the proposed algorithmic scheme. We note as well that model improvements are not restricted to the framework of energy systems and can inform researchers working on integrated planning problems in other application areas, including water resources (Naderi and Pishvaee 2017), industrial processes (Shah and Ierapetritou 2012), and facility location (An 2020).

1. *Superior Performance:* Our decomposed model consistently outperformed monolithic solution approaches using state-of-the-art commercial solvers on cases with discrete, integer investment decisions and was competitive with linearized investment decisions depending on model size and structure tested. Our ability to decouple and parallelize subperiods and provide $|W|$ cuts per iteration decreased both the number of iterations and runtime per iteration relative to a more conventional Benders decomposition algorithm, resulting in linear runtime increase as a function of resolution.

2. *Increased Resolution:* Because runtime scales well with master model size, a great deal of information on investments (including geographic constraints for expansion, integer transmission expansion, and different constraints or properties for different units of thermal plant) can be included with minimal expense to performance. Our decomposition algorithm's runtime scales linearly with number of weeks included, which further allows for inclusion of more subperiods beyond the standard strategy of considering only few weeks or days of operation without risking intractability. Decreasing the level of abstraction in this way helps eliminate structural uncertainty (Pfenninger 2017). Indeed, we demonstrated that increasing temporal resolution decreased the MSE associated with investment decisions by resource type and location.

5.1. Future Directions and Novel Capabilities

While increased resolution (as a corollary, decreased reliance on abstraction) is one exciting implication of this work, it is not the only application of the proposed algorithmic scheme. Improved performance also benefits research due to the additional analysis enabled by decreased computational bandwidth. Some potential additional applications include:

1. *Enabling More Extensive Analysis:* Improved computational performance due to decomposition can also be employed to conduct more extensive exploration of parameter uncertainty via scenario analysis methods or enable incorporation of methods like multiobjective optimization or modeling to generate alternatives (Patankar and Jenkins 2020), which explores a wide range of alternative solutions with similar objective function results (e.g., costs). Such methods require solving the planning problem multiple times under different parameters or objective functions, thus benefiting greatly from improved runtime for each problem.

2. *Ability to Capture Economies of Scale:* Making integer investments tractable permits modeling of discrete capacity investment decisions that capture economics of unit scale that are common in many applications, including electricity transmission and generator investment decisions. Other methods with continuous capacity decisions cannot capture economies of unit scale and may result in unrealistic biases (Donohoo-Vallett 2014). Additionally, without decomposition, modeling electricity network and power flows with KVL introduces nonlinearities in combined investment/operations models and generally requires Big-M reformulations that can significantly slow computational time loosening the convex relaxation of the MILP problem. The structure of the decomposition algorithm herein separates investment and operations into discrete problems and thus can allow KVL and other operational characteristics that have interactions with investments to be included seamlessly in a convex model.

3. *Model Accessibility.* Current state-of-the-art energy systems planning tools are inaccessible to many potential users due to high computational demands and their associated need for expensive commercial solvers. In fact, the benchmarks computed by Hans Mittelmann (Mittelmann 2023) report that the fastest open source solver can still be, on average, 30 to 40 times slower than the best commercial solvers. Analogously, Han et al. (2021) show that solvers that are completely open source (score "10" on the "openness" scale) tend to perform poorly on large-scale security constrained economic dispatch models. The decomposition method introduced here involves solving substantially smaller master and operational subproblems, which makes it easier to implement open-source solvers, thereby increasing accessibility of macro-energy systems planning models.

References

An K (2020) Battery electric bus infrastructure planning under demand uncertainty. *Transportation Res. Part C Emerg. Tech.* 111:572–587.

Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: A fresh approach to numerical computing. *SIAM Rev.* 59(1):65–98.

Bistline J, Abhyankar N, Blanford G, Clarke L, Fakhry R, McJeon H, Reilly J, et al. (2022) Actions for reducing us emissions at least 50% by 2030. *Science* 376(6596):922–924.

Brown T, Schlachtberger D, Kies A, Schramm S, Greiner M (2018) Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. *Energy* 160:720–739.

Cho S, Li C, Grossmann IE (2022) Recent advances and challenges in optimization models for expansion planning of power systems and reliability optimization. *Comput. Chemical Engrg.* 165:107924.

Donohoo-Vallett PE (2014) Design of wide-area electric transmission networks under uncertainty: Methods for dimensionality reduction. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.

Dunning I, Huchette J, Lubin M (2017) JuMP: A modeling language for mathematical optimization. *SIAM Rev.* 59(2):295–320.

Frew BA, Jacobson MZ (2016) Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the power model. *Energy* 117:198–213.

Frysztacki MM, Recht G, Brown T (2022) A comparison of clustering methods for the spatial reduction of renewable electricity optimisation models of Europe. *Energy Inform.* 5(1):4.

Han P, Xiong X, Zeng D, Feng Y, Zhao L, Yan Z (2021) Comprehensive evaluation of various linear programming solvers for SCED in power systems. *2021 IEEE 5th Conf. Energy Internet Energy System Integration (EI2)* (IEEE, Piscataway, NJ), 1058–1062.

He G, Mallapragada DS, Bose A, Heuberger-Austin CF, Gençer E (2021) Sector coupling via hydrogen to lower the cost of energy system decarbonization. *Energy Environ. Sci.* 14(9):4635–4646.

Helistö N, Kiviluoma J, Morales-España G, O'Dwyer C (2021) Impact of operational details and temporal representations on investment planning in energy systems dominated by wind and solar. *Appl. Energy* 290:116712.

Lara CL, Mallapragada DS, Papageorgiou DJ, Venkatesh A, Grossmann IE (2018) Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. *Eur. J. Oper. Res.* 271(3):1037–1054.

Larson E, Greig C, Jenkins J, Mayfield E, Pascale A, Zhang C, Drossman J, et al. (2021) Net zero America: Potential pathways, infrastructure, and impacts. Final report summary, Princeton University, Princeton, NJ.

Li C, Conejo AJ, Liu P, Omell BP, Siirola JD, Grossmann IE (2022) Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems. *Eur. J. Oper. Res.* 297(3):1071–1082.

Lohmann T, Rebennack S (2017) Tailored benders decomposition for a long-term power expansion model with short-term demand response. *Management Sci.* 63(6):2027–2048.

Mallapragada DS, Sepulveda NA, Jenkins JD (2020) Long-run system value of battery energy storage in future grids with increasing wind and solar generation. *Appl. Energy* 275:115390.

Mallapragada DS, Papageorgiou DJ, Venkatesh A, Lara CL, Grossmann IE (2018) Impact of model resolution on scenario outcomes for electricity sector system expansion. *Energy* 163:1231–1244.

Mittelmann H (2023) Benchmarks for optimization software. Accessed May 1, 2023, http://plato.asu.edu/bench.html.

Munoz FD, Hobbs BF, Watson JP (2016) New bounding and decomposition approaches for MILP investment problems: Multi-area transmission and generation planning under policy constraints. *Eur. J. Oper. Res.* 248(3):888–898.
Naderi MJ, Pishvaee MS (2017) A stochastic programming approach to integrated water supply and wastewater collection network design problem. *Comput. Chemical Engrg.* 104:107–127.
Neumann F, Hagenmeyer V, Brown T (2022) Assessments of linear power flow and transmission loss approximations in coordinated capacity expansion problems. *Appl. Energy* 314:118859.
Palmintier BS, Webster MD (2013) Heterogeneous unit clustering for efficient operational flexibility modeling. *IEEE Trans. Power Syst.* 29(3):1089–1098.
Palmintier BS, Webster MD (2015) Impact of operational flexibility on electricity generation planning with renewable and carbon targets. *IEEE Trans. Sustain. Energy.* 7(2):672–684.
Patankar N, Jenkins J (2020) Land use, transmission expansion, and supply chain scale-up implications of alternative scenarios for 100% carbon-free electricity generation in the American west. AGU Fall Meeting Abstracts, 2020:GC080-08.
Pfenninger S (2017) Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. *Appl. Energy* 197:1–13.
Poncelet K, Delarue E, D'haeseleer W (2020) Unit commitment constraints in long-term planning models: Relevance, pitfalls and the role of assumptions on flexibility. *Appl. Energy* 258:113843.
Poncelet K, Delarue E, Six D, Duerinck J, D'haeseleer W (2016) Impact of the level of temporal and operational detail in energy-system planning models. *Appl. Energy* 162:631–643.
Ricks W, Norbeck J, Jenkins J (2022) The value of in-reservoir energy storage for flexible dispatch of geothermal power. *Appl. Energy* 313:118807.
Ricks W, Xu Q, Jenkins JD (2023) Minimizing emissions from grid-based hydrogen production in the United States. *Environ. Res. Lett.* 18(1):014025.
Ringkjøb HK, Haugan PM, Solbrekke IM (2018) A review of modelling tools for energy and electricity systems with large shares of variable renewables. *Renew. Sustain. Energy Rev.* 96:440–459.
Schivley G, Welty E, Patankar N (2021) Powergenome/powergenome: v0.4.0 (v0.4.0). Zenodo. https://doi.org/10.5281/zenodo.4426097.
Sepulveda NA (2020) *Decarbonization of power systems, multi-stage decision-making with policy and technology uncertainty*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.
Shah NK, Ierapetritou MG (2012) Integrated production planning and scheduling optimization of multisite, multiproduct process industry. *Comput. Chemical Engrg.* 37:214–226.
Siala K, Mahfouz MY (2019) Impact of the choice of regions on energy system models. *Energy Strategy Rev.* 25:75–85.
Victoria M, Zeyen E, Brown T (2022) Speed of technological transformations required in Europe to achieve different climate goals. Joule 6(5):1066–1086.
Victoria M, Zhu K, Brown T, Andresen GB, Greiner M (2020) The role of photovoltaics in a sustainable European energy system under variable CO_2 emissions targets, transmission capacities, and costs assumptions. *Prog Photovoltaics Res. Appl.* 28(6):483–492.
Xu Q, Hobbs BF (2019) Value of model enhancements: Quantifying the benefit of improved transmission planning models. *IET Gener. Transm. Distrib.* 13(13):2836–2845.