Nested Benders's Decomposition of Capacity-Planning Problems for Electricity Systems with Hydroelectric and Renewable Generation

Kenjiro Yagi[1] and Ramteen Sioshansi[2,3,4,5,6*]

[1]Tokushu Tokai Paper Co., Ltd., Tokyo, 104-0028, Japan.
[2*]Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, 15213-3815, Pennsylvania, United States of America.
[3]Carnegie Mellon Electricity Industry Center, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, 15213-3815, Pennsylvania, United States of America.
[4]Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, 15213-3815, Pennsylvania, United States of America.
[4]Wilton E. Scott Institute for Energy Innovation, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, 15213-3815, Pennsylvania, United States of America.
[6]Department of Integrated Systems Engineering, The Ohio State University, 1971 Neil Avenue, Columbus, 43210-1271, Ohio, United States of America.

*Corresponding author(s). E-mail(s): rsioshan@andrew.cmu.edu;
Contributing authors: yagi.7@buckeyemail.osu.edu;

Abstract

Nested Benders's decomposition is an efficient means to solve large-scale optimization problems with a natural time sequence of decisions. This paper examines the use of the technique to decompose and solve efficiently capacity-expansion problems for electricity systems with hydroelectric and renewable generators. To this end we develop an archetypal planning model that captures key features of hydroelectric and renewable generators and apply it to a case study that is based on the Columbia River system in the northwestern United States of

America. We apply standard network and within-year temporal simplifications to reduce the problem's size. Nevertheless, the remaining problem is large-scale and we demonstrate the use of nested Benders's decomposition to solve it. We explore refinements of the decomposition method which yield further performance improvements. Overall, we show that nested Benders's decomposition yields good computational performance with minimal loss of model fidelity.

Keywords: Electricity-system planning, decomposition, hydroelectric generation, renewable generation

1 Introduction

Capacity-planning problems for electricity systems can be large-scale and computationally challenging. As such, the literature explores approaches to make these problems more tractable. One standard approach is to simplify the representation of system operations, for instance by reducing the number of operating periods that are modeled. A classic reduction approach uses the electricity system's load-duration curve to approximate its operations (Stoft, 2002; Sioshansi, 2016). Another reduction approach uses representative time periods.

Representative time periods raise questions regarding how to select them and their durations. One approach to selecting representative time periods is with an heuristic. For instance, one can select representative time periods based on season and peak-load conditions (Cohen et al, 2019). Another selection approach uses data features, *e.g.*, clustering (Nahmmacher et al, 2016; Liu et al, 2018a; Boffino et al, 2019). A third approach is to treat the selection problem as its own optimization (Poncelet et al, 2017).

The duration of the representative operating periods can be especially important if the flexibility needs of an electricity system increases, *e.g.*, due to variable renewable-energy sources. In such a case, it can be important to capture intertemporal constraints, *e.g.*, ramping limits. This modeling need may call for using representative operating periods, *e.g.*, days or weeks, that can capture such constraints. Liu et al (2018b) use 30 days to represent electricity-system operations over the course of each year in a capacity-planning model that they apply to a case study that is based on Texas. They demonstrate that including as opposed to relaxing intraday ramping constraints in their model yields different generation-capacity mixes. Maluenda et al (2018) use four representative days for modeling the Chilean electricity system. They show that this approach represents electricity-system operations better than a load-block model does, because the latter does not capture inter-hour operating constraints.

Many works examine simplifications of capacity-planning problems for electricity systems with thermal and renewable-energy resources. There is a much more limited literature that examines such simplifications for electricity systems with hydroelectric resources, which yield additional complexities in capacity planning. Hydroelectric generators that are on a cascaded river system have interdependencies and coupling

constraints that complicate their operations. As the flexibility needs of an electricity system increases, it becomes more important to capture accurately the flexibility of hydroelectric resources and the impacts of cascaded river systems thereupon (Huertas-Hernando et al, 2017). As such, investment decisions for electricity systems that have heavy dependence upon hydroelectric resources can be influenced by the fidelity with which hourly operations of hydroelectric resources are modeled. Ibanez et al (2014) find that modeling the details of the operations of a cascaded river yields better operational performance, *e.g.*, less renewable-energy curtailment and operational cost.

A second complication that hydroelectric generators raise is that some have large reservoirs with seasonal or inter-annual water-storage capacities. Natural water inflows into such reservoirs (or into cascaded river systems) can have seasonal and inter-annual variability and uncertainty. As such, capacity planning of an electricity system with hydroelectric resources may require capturing conditions over many years. In addition, the water-storage capabilities of hydroelectric reservoirs means that the chronology of hydroelectric operations should be captured over long periods of time, *e.g.*, a year or longer.

Given these complexities of capacity planning of electricity systems with high penetrations of hydroelectric generation, this and a companion paper (Yagi and Sioshansi, 2023) explore different aspects of simplifying such problems. Yagi and Sioshansi (2023) focus on model-reduction approaches, including the selection of representative operating periods, which yield problems that are smaller than an unreduced model. Such a reduced model can remain computationally taxing, which motivates the current paper, which examines the use of nested Benders's decomposition (Benders, 1962; Birge, 1985) to make the resultant problem more computationally tractable.

Our use of Benders's decomposition is not novel, as is evidenced by the literature survey of Rahmaniani et al (2017). Benders's decomposition has a rich history of application to the electricity sector, especially in the presence of hydroelectric resources, which can involve temporally coupled decisions. The seminal work of Pereira and Pinto (1985) employs a stochastic variant of Benders's decomposition for operational and water-use optimization of a cascaded hydroelectric system. Hjelmeland et al (2019) model optimal medium-term operation of hydroelectric generation. Contrasting these two works shows the benefits of improved hardware and optimization software during the intervening years. Pereira and Pinto (1985) optimize water use over a five-month optimization horizon with monthly temporal resolution, whereas Hjelmeland et al (2019) examine weekly water use over a two-year planning horizon. Maceiral et al (2018) provide a comprehensive survey of the successful use of Benders's decomposition for optimizing hydroelectric-resource operation. The value of Benders's decomposition is evidenced further by Chile's energy law, which requires its use to optimize the operation of the nation's hydroelectric resources (Raineri, 2006).

To date, Benders's decomposition sees limited application to capacity-planning problems for electricity systems that include hydroelectric resources. Falugi et al (2016) apply nested Benders's decomposition to optimize transmission-infrastructure planning for an electricity system. Lara (2019) employs nested Benders's decomposition to simplify capacity-expansion models for electricity systems. However, neither of these

works has considerable focus on the complexities that are inherent with the inclusion of hydroelectric generation in an electricity system.

Thus, the novelty and primary contribution of our work is to demonstrate that nested Benders's decomposition can yield a computationally tractable and high-fidelity capacity-planning model for an electricity system with high penetrations of renewable and hydroelectric generation. We find that the standard 'textbook' implementation of Benders's decomposition can yield computational challenges and demonstrate simple techniques to overcome these. Furthermore, we demonstrate that the benefit of Benders's decomposition can be exploited and strengthened if the reduction methods that are explored by Yagi and Sioshansi (2023) are employed. In particular, we show that the representative operating periods that are studied by Yagi and Sioshansi (2023) can be combined effectively with nested Benders's decomposition. We use economic regret as our primary solution-quality metric (Ramírez-Sagner and Muñoz, 2019). Many works (Merrick, 2016; Liu et al, 2018b,a; Boffino et al, 2019; Barrera-Santana and Sioshansi, 2023) compare the optimized objective-function value to assess solution quality. We find that economic regret provides a more robust solution-quality measure (Yagi and Sioshansi, 2023). Taken together, this paper and the work of Yagi and Sioshansi (2023) provide complementary approaches to simplify capacity planning for electricity systems with high penetrations of renewable and hydroelectric resources.

We proceed through the remainder of the paper as follows. Section 2 provides the archetypal capacity-planning model, which is the basis of our analysis. We apply the techniques from our companion paper (Yagi and Sioshansi, 2023) to reduce the model. Section 3 outlines the decomposition technique, as well as some refinements to ease its implementation. Section 4 provides data for a comprehensive case study to which we apply our model and decomposition technique. The case study is based on the electricity system within Western Electricity Coordination Council (WECC), which is in western North America. WECC includes the Columbia River system, which is a major cascaded hydroelectric system. Section 5 provides case-study results and Sect. 6 concludes.

2 Model

This section provides the formulation of our archetypal capacity-expansion model. This model does not capture every nuance of electricity-system planning. Rather, the model captures key features of capacity planning with hydroelectric, renewable, and thermal resources. Moreover, this model employs some of the reductions that are examined in our companion paper (Yagi and Sioshansi, 2023). The companion paper provides a more generic capacity-expansion model and detailed evaluations of the reductions that we employ here.

2.1 Model Notation

We begin by defining the following sets and indices.

c	index for hydroelectric dams
g	index for non-hydroelectric-generation technologies
$\mathcal{G}$	set of non-hydroelectric-generation technologies

h	index for hours during a year
$\mathcal{H}$	set of hydroelectric dams
$\bar{\mathcal{H}}_c$	set of dams that are upstream of dam c
$\mathcal{H}_n^{\mathcal{L}}$	set of dams that are at transmission node n
l	index for transmission lines
$\mathcal{L}$	set of transmission lines
$\mathcal{L}_n^{\text{in}}$	set of transmission lines that flow into node n in the nominal direction
$\mathcal{L}_n^{\text{out}}$	set of transmission lines that flow out from node n in the nominal direction
n	index for transmission nodes
$\mathcal{N}$	set of transmission nodes
$\mathcal{T}^{\text{H}}$	ordered set of hours during a year
y	index for years
$\mathcal{Y}$	ordered set of years
z	index of pieces of approximation of transmission losses
$\mathcal{Z}$	set of pieces of approximation of transmission losses

Next, we define the following parameters and functions.

$C_{y,g,n}^{\mathcal{G},\text{i}}$	annualized year-y capital cost of non-hydroelectric-generation technology g that is located at node n (\$/MW)
$C_{y,g,n,h}^{\mathcal{G},\text{V}}$	operation cost during hour h of year y of non-hydroelectric-generation technology g that is located at node n (\$/MWh)
$C^{\mathcal{L}}$	annualized capital cost of transmission-line capacity (\$/MW)
C^{U}	value of lost load (\$/MWh)
D_z	coefficient for zth piece of flow-related transmission losses (p.u.)
E_z	coefficient for zth piece of capacity-related transmission losses (p.u.)
$k_{0,g,n}^{\mathcal{G},\text{i}}$	initial capacity that is present at node n of non-hydroelectric-generation technology g (MW)
$k_{0,l}^{\mathcal{L}}$	initial capacity that is present of transmission line l (MW)
$L_{y,n,h}$	node-n electric load during hour h of year y (MW)
$Q_c^{\mathcal{H},+}$	maximum output of generator at dam c (MW)
R_c	ramp rate of generator at dam c (p.u.)
R_g	ramp rate of non-hydroelectric-generation technology g (p.u.)
$\bar{W}_c^{\mathcal{H}}$	maximum water level of dam c (acre-feet)
$W_c^{\mathcal{H},-}$	minimum water flow from dam c (acre-feet)
$W_c^{\mathcal{H},+}$	maximum water flow from dam c (acre-feet)
$W_{y,c,h}^{\mathcal{H},f}$	natural water inflow to dam c during hour h of year y (acre-feet)
$w_{1,c,0}^{\mathcal{H},\text{str}}$	initial water level of dam c as of the beginning of the planning horizon (acre-feet)
$\beta_{c,0}$	constant in the function that models power output of the generator that is located at dam c (MW)
$\beta_{c,1}$	coefficient that multiplies the water flow through the generator in the function that models power output of the generator that is located at dam c (MW/acre-feet)
$\beta_{c,2}$	coefficient that multiplies the water level of its reservoir in the function that models power output of the generator that is located at dam c (MW/acre-feet)

Γ discount rate (p.u.)
ι^{+} maximum capacity that can be added from one year to the next (p.u.)
ι_{g}^{-} maximum capacity of non-hydroelectric-generation technology g that can be retired from one year to the next (p.u.)
κ_{g}^{E} output-based CO_2-equivalent-emission rate of non-hydroelectric-generation technology g (t/MWh)
κ^{T} tax rate for CO_2-equivalent emissions (\$/t)
ρ cost of retiring non-hydroelectric-generating capacity (p.u.)
$\sigma_{y,g}$ minimum total energy supply from non-hydroelectric-generation technology g during year y (p.u.)
$\Upsilon_{y,h}$ weight on hour h of year y (h)
$\Phi_{y,g,n,h}$ availability factor during hour h of year y for non-hydroelectric-generation technology g that is located at node n (p.u.)

Finally, we define the following set of decision variables.

$f_{y,l,h}^{\mathcal{L}}$ net power flow during hour h of year y through transmission line l (MW)
$k_{y,g,n}^{\mathcal{G},\mathrm{i}}$ total capacity during year y of non-hydroelectric-generation technology g that is present at node n (MW)
$k_{y,g,n}^{\mathcal{G},\rho}$ capacity of non-hydroelectric-generation technology g that is retired during year y from node n (MW)
$k_{y,l}^{\mathcal{L}}$ total capacity during year y of transmission line l (MW)
$q_{y,g,n,h}^{\mathcal{G}}$ power output during hour h of year y of non-hydroelectric-generation technology g that is located at node n (MW)
$q_{y,c,h}^{\mathcal{H}}$ power output during hour h of year y of generator that is located at dam c (MW)
$q_{y,n,h}^{\mathrm{U}}$ unserved node-n load during hour h of year y (MW)
$w_{y,c,h}^{\mathcal{H},\mathrm{g}}$ water flow during hour h of year y through generator that is located at dam c (acre-feet)
$w_{y,c,h}^{\mathcal{H},\mathrm{spl}}$ water spilled during hour h of year y from dam c (acre-feet)
$w_{y,c,h}^{\mathcal{H},\mathrm{str}}$ water level as of the end of hour h of year y of dam c (acre-feet)
$\mu_{y,l,h}^{\mathcal{L}}$ power losses during hour h of year y on transmission line l (MW)

2.2 Model Formulation

The capacity-planning model is formulated as:

$$\min \sum_{y\in\mathcal{Y}} \Gamma^{y-1}\cdot\left\{\sum_{l\in\mathcal{L}} C^{\mathcal{L}}k_{y,l}^{\mathcal{L}} + \sum_{n\in\mathcal{N}}\left\{\sum_{g\in\mathcal{G}}\left[C_{y,g,n}^{\mathcal{G},\mathrm{i}}\cdot\left(k_{y,g,n}^{\mathcal{G},\mathrm{i}} + \rho k_{y,g,n}^{\mathcal{G},\rho}\right) + \sum_{h\in\mathcal{T}^{\mathrm{H}}}\Upsilon_{y,h}\cdot\left(C_{y,g,n,h}^{\mathcal{G},\mathrm{V}} + \kappa^{\mathrm{T}}\kappa_{g}^{\mathrm{E}}\right)q_{y,g,n,h}^{\mathcal{G}}\right] + \sum_{h\in\mathcal{T}^{\mathrm{H}}}\Upsilon_{y,h}C^{\mathrm{U}}q_{y,n,h}^{\mathrm{U}}\right\}\right\} \tag{1}$$

$$\text{s.t. } k_{y,g,n}^{\mathcal{G},\rho} \geq k_{y-1,g,n}^{\mathcal{G},\mathrm{i}} - k_{y,g,n}^{\mathcal{G},\mathrm{i}}; \forall y\in\mathcal{Y}, g\in\mathcal{G}, n\in\mathcal{N} \tag{2}$$

$$k^{\mathcal{G},\mathrm{i}}_{y-1,g,n} - \iota^{-}_{g} k^{\mathcal{G},\mathrm{i}}_{0,g,n} \leq k^{\mathcal{G},\mathrm{i}}_{y,g,n} \leq (1+\iota^{+}) k^{\mathcal{G},\mathrm{i}}_{y-1,g,n};$$
$$\forall y \in \mathcal{Y}, g \in \mathcal{G}, n \in \mathcal{N} \tag{3}$$
$$k^{\mathcal{L}}_{y-1,l} \leq k^{\mathcal{L}}_{y,l} \leq (1+\iota^{+}) k^{\mathcal{L}}_{y-1,l}; \forall y \in \mathcal{Y}, l \in \mathcal{L} \tag{4}$$
$$\sum_{n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}}} \Upsilon_{y,h} q^{\mathcal{G}}_{y,g,n,h} \geq \sigma_{y,g} \sum_{n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}}} \Upsilon_{y,h} L_{y,n,h}; \forall y \in \mathcal{Y}, g \in \mathcal{G} \tag{5}$$
$$\sum_{g \in \mathcal{G}} q^{\mathcal{G}}_{y,g,n,h} + \sum_{c \in \mathcal{H}^{\mathcal{L}}_{n}} q^{\mathcal{H}}_{y,c,h} + \sum_{l \in \mathcal{L}^{\mathrm{in}}_{n}} \left(f^{\mathcal{L}}_{y,l,h} - \frac{1}{2} \mu^{\mathcal{L}}_{y,l,h} \right) + q^{\mathrm{U}}_{y,n,h}$$
$$- \sum_{l \in \mathcal{L}^{\mathrm{out}}_{n}} \left(f^{\mathcal{L}}_{y,l,h} + \frac{1}{2} \mu^{\mathcal{L}}_{y,l,h} \right) = L_{y,n,h}; \forall y \in \mathcal{Y}, n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}} \tag{6}$$
$$0 \leq q^{\mathcal{G}}_{y,g,n,h} \leq \Phi_{y,g,n,h} k^{\mathcal{G},\mathrm{i}}_{y,g,n}; \forall y \in \mathcal{Y}, g \in \mathcal{G}, n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}} \tag{7}$$
$$- R_g k^{\mathcal{G},\mathrm{i}}_{y,g,n} \leq q^{\mathcal{G}}_{y,g,n,h} - q^{\mathcal{G}}_{y,g,n,h-1} \leq R_g k^{\mathcal{G},\mathrm{i}}_{y,g,n};$$
$$\forall y \in \mathcal{Y}, g \in \mathcal{G}, n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}} \tag{8}$$
$$q^{\mathcal{G}}_{y,g,n,0} = q^{\mathcal{G}}_{y-1,g,n,|\mathcal{T}^{\mathrm{H}}|}; \forall y \in \mathcal{Y}, y \neq 1, g \in \mathcal{G}, n \in \mathcal{N} \tag{9}$$
$$0 \leq q^{\mathrm{U}}_{y,n,h} \leq L_{y,n,h}; \forall y \in \mathcal{Y}; n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}} \tag{10}$$
$$q^{\mathcal{H}}_{y,c,h} = \beta_{c,0} + \beta_{c,1} w^{\mathcal{H},\mathrm{g}}_{y,c,h} + \beta_{c,2} w^{\mathcal{H},\mathrm{str}}_{y,c,h}; \forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}} \tag{11}$$
$$- R_c Q^{\mathcal{H},+}_{c} \leq q^{\mathcal{H}}_{y,c,h} - q^{\mathcal{H}}_{y,c,h-1} \leq R_c Q^{\mathcal{H},+}_{c};$$
$$\forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}} \tag{12}$$
$$q^{\mathcal{H}}_{y,c,0} = q^{\mathcal{H}}_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}; \forall y \in \mathcal{Y}, y \neq 1, c \in \mathcal{H} \tag{13}$$
$$w^{\mathcal{H},\mathrm{str}}_{y,c,h} = w^{\mathcal{H},\mathrm{str}}_{y,c,h-1} + \Upsilon_{y,h} \cdot \left(W^{\mathcal{H},f}_{y,c,h} - w^{\mathcal{H},\mathrm{g}}_{y,c,h} - w^{\mathcal{H},\mathrm{spl}}_{y,c,h} \right)$$
$$+ \sum_{c' \in \bar{\mathcal{H}}_{c}} \Upsilon_{y,h} \cdot \left(w^{\mathcal{H},\mathrm{g}}_{y,c',h} + w^{\mathcal{H},\mathrm{spl}}_{y,c',h} \right); \forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}} \tag{14}$$
$$w^{\mathcal{H},\mathrm{str}}_{y,c,0} = w^{\mathcal{H},\mathrm{str}}_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}; \forall y \in \mathcal{Y}, y \neq 1, c \in \mathcal{H} \tag{15}$$
$$0 \leq w^{\mathcal{H},\mathrm{str}}_{y,c,h} \leq \bar{W}^{\mathcal{H}}_{c}; \forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}} \tag{16}$$
$$W^{\mathcal{H},-}_{c} \leq w^{\mathcal{H},\mathrm{g}}_{y,c,h} + w^{\mathcal{H},\mathrm{spl}}_{y,c,h} \leq W^{\mathcal{H},+}_{c}; \forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}} \tag{17}$$
$$w^{\mathcal{H},\mathrm{str}}_{|\mathcal{Y}|,c,|\mathcal{T}^{\mathrm{H}}|} \geq w^{\mathcal{H},\mathrm{str}}_{1,c,0}; \forall c \in \mathcal{H} \tag{18}$$
$$- k^{\mathcal{L}}_{y,l} \leq f^{\mathcal{L}}_{y,l,h} \leq k^{\mathcal{L}}_{y,l}; \forall y \in \mathcal{Y}, l \in \mathcal{L}, h \in \mathcal{T}^{\mathrm{H}} \tag{19}$$
$$\mu^{\mathcal{L}}_{y,l,h} \geq E_z k^{\mathcal{L}}_{y,l} + D_z f^{\mathcal{L}}_{y,l,h}; \forall y \in \mathcal{Y}, l \in \mathcal{L}, h \in \mathcal{T}^{\mathrm{H}}, z \in \mathcal{Z} \tag{20}$$
$$\mu^{\mathcal{L}}_{y,l,h} \geq E_z k^{\mathcal{L}}_{y,l} - D_z f^{\mathcal{L}}_{y,l,h}; \forall y \in \mathcal{Y}, l \in \mathcal{L}, h \in \mathcal{T}^{\mathrm{H}}, z \in \mathcal{Z} \tag{21}$$
$$k^{\mathcal{G},\rho}_{y,g,n} \geq 0; \forall y \in \mathcal{Y}, g \in \mathcal{G}, n \in \mathcal{N} \tag{22}$$
$$w^{\mathcal{H},\mathrm{g}}_{y,c,h}, w^{\mathcal{H},\mathrm{spl}}_{y,c,h} \geq 0; \forall y \in \mathcal{Y}, c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}}. \tag{23}$$

We represent time in the model using two different scales. Capacity-addition and -retirement decisions are made annually, over the ordered set, $\mathcal{Y}$, of years. Between these

annual investment and retirement epochs, electricity-system operations are captured during an ordered set, $\mathcal{T}^{\mathrm{H}}$, of representative operating hours. These operating hours are ordered to capture the chronology of decisions and parameters, which affect water levels in the hydroelectric system. The temporal granularity of the decisions is selected arbitrarily without loss of model generality.

Objective function (1) minimizes the discounted sum of capacity and operational costs over this time horizon and contains four cost terms. The first represents the cost of adding transmission capacity. For simplicity, our model does not allow for transmission retirements. The second term represents the cost of adding or retiring generation capacity. The capacity of the hydroelectric system is fixed. However, the capacities of other generation technologies can be adjusted. The third term is the cost of operating the generating fleet. Our model accommodates a price on CO_2-equivalent emissions, *e.g.*, as a carbon-policy mechanism (Liu et al, 2021; Yagi and Sioshansi, 2021). Operating the hydroelectric system is assumed to have no direct cost. The final term in (1) gives the load-curtailment cost.

Constraints (2)–(4) restrict capacity-related decisions. Constraint set (2) defines the amount of generation capacity that is retired during each year by the decrease in generation capacity between that year and the preceding one. Constraint sets (3) and (4) impose limits on capacity changes from one year to the next (*e.g.*, resource or budget limits).

The remaining constraints pertain to electricity-system operations. Constraint set (5) can be used to model a renewable-portfolio standard, which is one approach to drive renewable-energy adoption (Sioshansi, 2016). Constraint set (5) imposes a minimum-energy requirement, whereas some renewable-portfolio standards are based on installed capacity. The two types of standards are similar, insomuch as one may account for the capacity factor of the renewable-energy technology in setting the standard. Constraint set (6) imposes hourly load balance between energy that is supplied and consumed at each electricity-system node. Constraint sets (7) and (8) impose minimum and maximum capacities and interhourly ramping limits, respectively, on power output from non-hydroelectric generators. The parameter, $\Phi_{y,g,n,h}$, which appears in the right-hand side of (7) can capture the variability of weather-dependent renewable generators. Constraint set (9) defines the starting output level of each non-hydroelectric generator as of the beginning of each year in terms of its ending output level as of the end of the previous year. Constraint set (10) limits curtailed load to be no greater than demand.

Constraint set (11) relates the power output of each hydroelectric generator to water flow through the generator and the water level of the dam. This linear relationship yields a reasonable trade-off between model fidelity and tractability (Yagi and Sioshansi, 2023). More complex non-linear or non-convex relationships appear in the literature (Hidalgo et al, 2014; Hunter-Rinderle and Sioshansi, 2021). Although such relationships could be substituted in-place of (11), a linear model structure is needed for application of nested Benders's decomposition. Constraint set (12) imposes ramping limits on the hydroelectric generators. Constraint set (13) defines the starting output level of each hydroelectric generator as of the beginning of each year in terms of its ending output level as of the end of the previous year. Constraint set (14) enforces

hourly water balance for each dam. The ending hour-h water level of a dam is defined as the sum of its ending hour-$(h-1)$ water level and inflows from upstream dams and natural inflows, less the sum of water that is released through the generator and for spillage. We assume zero water-travel times between upstream and downstream dams. Our companion paper (Yagi and Sioshansi, 2023) shows that this assumption has negligible model-fidelity effects. Constraint set (15) defines the starting water level of each dam as of the beginning of each year in terms of its ending water level from the previous year. Constraint set (16) imposes minimum and maximum water-level limits on each dam. Constraint set (17) imposes minimum and maximum limits on water flows through each dam. These constraints could represent physical limits on a dam or environmental considerations, *e.g.*, needed water levels for wildlife preservation. Constraint set (18) forces the ending water level of each dam to be no less than its initial level. Such a constraint is an heuristic approach to ascribe value to carrying water as of the end of the model horizon (Sioshansi et al, 2022).

Constraint set (19) imposes power-flow limits on transmission lines. Constraint sets (20) and (21) define transmission losses on each line as a piecewise-linear function of line capacity and flow. We model the transmission system assuming that flows can be directed within the network. Our approach to modeling transmission losses can provide a reasonable approximation of a linearization of Kirchhoff's circuit laws (Ahlhaus and Stursberg, 2013).

Constraints (22) and (23) impose non-negativity on generation-capacity retirements and water flows.

3 Nested Benders's Decomposition

Our implementation of Nested Benders's decomposition exploits the sequential time structure of (1)–(23) and of the decisions that are optimized therein. We decompose (1)–(23) by dividing it into $|\mathcal{Y}|$ subproblems, each of which corresponds to a $y \in \mathcal{Y}$.

To detail our decomposition approach, we begin by defining, $\forall y \in \mathcal{Y}$, a_y as a vector of all of the year-y decision variables. Thus, $\forall y \in \mathcal{Y}$, a_y consists of $f^{\mathcal{L}}_{y,l,h}$ (transmission-line power flows) and $\mu^{\mathcal{L}}_{y,l,h}$ (transmission-lines losses), $\forall l \in \mathcal{L}, h \in \mathcal{T}^{\mathrm{H}}$; $k^{\mathcal{G},\mathrm{i}}_{y,g,n}$ (non-hydroelectric-generation capacities) and $k^{\mathcal{G},\rho}_{y,g,n}$ (retired non-hydroelectric-generation capacities), $\forall g \in \mathcal{G}, n \in \mathcal{N}$; $k^{\mathcal{L}}_{y,l}$ (transmission-line capacities), $\forall l \in \mathcal{L}$; $q^{\mathcal{G}}_{y,g,n,h}$ (non-hydroelectric-generator outputs), $\forall g \in \mathcal{G}, n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}}$; $q^{\mathcal{H}}_{y,c,h}$ (hydroelectric-generator outputs), $w^{\mathcal{H},\mathrm{g}}_{y,c,h}$ (hydroelectric-generator water flows), $w^{\mathcal{H},\mathrm{spl}}_{y,c,h}$ (hydroelectric-dam water spillages), and $w^{\mathcal{H},\mathrm{str}}_{y,c,h}$ (hydroelectric-dam water levels), $\forall c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}}$; and $q^{\mathrm{U}}_{y,n,h}$ (curtailed electricity demands), $\forall n \in \mathcal{N}, h \in \mathcal{T}^{\mathrm{H}}$.

Next, $\forall y \in \mathcal{Y}$ we define s_y as a vector of year-y state variables. For $y = 1$, s_1 consists of $k^{\mathcal{G},\mathrm{i}}_{0,g,n}$ (ending year-0 non-hydroelectric-generation capacities), $\forall g \in \mathcal{G}, n \in \mathcal{N}$; $k^{\mathcal{L}}_{0,l}$ (ending year-0 transmission-line capacities), $\forall l \in \mathcal{L}$; and $w^{\mathcal{H},\mathrm{str}}_{1,c,0}$ (starting year-1 hydroelectric-dam water levels), $\forall c \in \mathcal{H}$. For all $y \in \mathcal{Y}$ such that $y > 1$, s_y consists of $k^{\mathcal{G},\mathrm{i}}_{y-1,g,n}$ (ending year-$(y-1)$ non-hydroelectric-generation capacities) and $q^{\mathcal{G}}_{y-1,g,n,|\mathcal{T}^{\mathrm{H}}|}$

(ending year-$(y-1)$ non-hydroelectric outputs), $\forall g \in \mathcal{G}, n \in \mathcal{N}$; $k^{\mathcal{L}}_{y-1,l}$ (ending year-$(y-1)$ transmission-line capacities), $\forall l \in \mathcal{L}$; and $q^{\mathcal{H}}_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}$ (ending year-$(y-1)$ hydroelectric-generator outputs) and $w^{\mathcal{H},\mathrm{str}}_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}$ (ending year-$(y-1)$ hydroelectric-dam water levels), $\forall c \in \mathcal{H}$. For all $y \in \mathcal{Y}$, s_y is a state variable in the sense that the value of s_y is needed to set starting conditions and specify constraints that restrict the choice of a_y. For all $y \in \mathcal{Y}$, s_{y+1} can be determined directly from a_y, *e.g.*, using (9), (13), and (15). Thus, for notational ease we denote the state variable as:

$$s_{y+1} = \eta(a_y); \tag{24}$$

for all $y \in \mathcal{Y}$, where $\eta(\cdot)$ is defined appropriately.

Next, $\forall y \in \mathcal{Y}$, we define $A_y(s_y)$ as the set of year-y variants of (2)–(8), (10)–(12), (14), (16), (17) and (19)–(23). In addition, we include (18) in the definition of $A_{|\mathcal{Y}|}(s_{|\mathcal{Y}|})$. Thus, $\forall y \in \mathcal{Y}$, the choice of a_y is constrained directly by $A_y(s_y)$. Constraint sets (9), (13), and (15) do not need to be included explicitly in defining $A_y(s_y)$, because the state variable, s_y, conveys the starting water level of each dam as of the beginning of year y. Next, $\forall y \in \mathcal{Y}$, we define:

$$\Psi_y(a_y) = \Gamma^{y-1} \cdot \left\{ \sum_{l \in \mathcal{L}} C^{\mathcal{L}} k^{\mathcal{L}}_{y,l} + \sum_{n \in \mathcal{N}} \left\{ \sum_{g \in \mathcal{G}} \left[C^{\mathcal{G},\mathrm{i}}_{y,g,n} \cdot \left(k^{\mathcal{G},\mathrm{i}}_{y,g,n} + \rho k^{\mathcal{G},\rho}_{y,g,n} \right) \right.\right.\right.$$
$$\left.\left.\left. + \sum_{h \in \mathcal{T}^{\mathrm{H}}} \Upsilon_{y,h} \cdot \left(C^{\mathcal{G},\mathrm{V}}_{y,g,n,h} + \kappa^{\mathrm{T}} \kappa^{\mathrm{E}}_{g} \right) q^{\mathcal{G}}_{y,g,n,h} \right] + \sum_{h \in \mathcal{T}^{\mathrm{H}}} \Upsilon_{y,h} C^{\mathrm{U}} q^{\mathrm{U}}_{y,n,h} \right\} \right\};$$

as the year-y objective function.

With these definitions, $\forall y \in \mathcal{Y}$, we can give the formulation of the following year-y subproblem:

$$\min_{a_y \in A_y(s_y), \theta_y \geq 0} \Psi_y(a_y) + \theta_y. \tag{25}$$

For all $y \in \mathcal{Y}$, the year-y subproblem takes as an input the year-y state variable, s_y, and determines an a_y that minimizes $\Psi_y(a_y)$. The auxiliary variable, θ_y, represents the resultant objective-function values of the decisions that are made subsequent to year y. More specifically, θ_y is meant to represent the optimized value of:

$$\sum_{y' \in \mathcal{Y}: y' > y} \Psi_{y'}(a_{y'}).$$

In other words, solving (25) for some $y \in \mathcal{Y}$ determines year-y decisions without direct consideration of previous or subsequent years. If there is a $y \in \mathcal{Y}$ such that θ^*_y solves (25) for y and:

$$\theta^*_y < \sum_{y' \in \mathcal{Y}: y' > y} \Psi_{y'}(a_{y'}); \tag{26}$$

then θ^*_y underestimates the true objective-function values of subsequent decisions and an optimality cut must be added to (25) for y. The optimality cut allows θ_y to provide

a better estimate of:

$$\sum_{y' \in \mathcal{Y}: y' > y} \Psi_{y'}(a_{y'}).$$

Otherwise, if (26) does not hold for any $y \in \mathcal{Y}$, then the solutions that are obtained from solving (25) $\forall y \in \mathcal{Y}$ are optimal in (1)–(23).

To separate these optimality cuts, we begin by defining $\forall y \in \mathcal{Y}$ such that $y > 1$, the linear dual of (25) as:

$$\max_{\lambda_y \in \Lambda_y(s_y)} \Pi_y(\lambda_y, s_y);$$

which, by (24), can be re-written as:

$$\max_{\lambda_y \in \Lambda_y(\eta(a_{y-1}))} \Pi_y(\lambda_y, \eta(a_{y-1})). \tag{27}$$

If for some $y \in \mathcal{Y}$, solving (25) yields θ_y^* that satisfies (26), then an optimality cut of the form:

$$\theta_y \geq \Pi_{y+1}(\lambda_{y+1}^*, \eta(a_y)); \tag{28}$$

where λ_{y+1}^* is optimal in year-$(y+1)$ linear dual (27), is added to (25). Benders (1962) provides the theoretical basis of the optimality cut, which arises from the strong-duality property of linear optimization problems.

A technical issue can arise in adding (28) to (25), because large-scale instances of (25) cannot be solved efficiently using Simplex method. Simplex method yields an optimal solution which is a basic feasible solution (Luenberger and Ye, 2008; Sioshansi and Conejo, 2017). A basic feasible solution has the property that non-basic variables are equal to zero. Large-scale instances of (25) must be solved using an interior-point or barrier method without crossover (Bertsekas, 1995). Such algorithms can yield solutions with near-zero dual-variable values, which would be zero if Simplex method is used instead. Optimality cuts with such near-zero coefficients can make (25) difficult to solve, due to numerical-scaling issues (*e.g.*, values of λ_{y+1}^* of that are near-zero but non-zero). As such, we employ a cutoff and fix equal to zero coefficients that are sufficiently small in magnitude.

Another issue that can arise in solving (25) is that for some $y \in \mathcal{Y}$, the state variable, s_y, could yield an infeasible problem. In our case, this can arise if the starting water level of a reservoir, $w_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}^{\mathcal{H},\mathrm{str}}$, is too low to satisfy some combination of (16)–(18) for some $c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}}$. One way to address this issue is by separating and adding feasibility cuts to (25), which we find yields poor computational performance (Yagi, 2020). An alternative approach is to add artificial slack variables, which ensure that (16)–(18) are feasible for all possible values of $w_{y-1,c,|\mathcal{T}^{\mathrm{H}}|}^{\mathcal{H},\mathrm{str}}$, $\forall c \in \mathcal{H}, h \in \mathcal{T}^{\mathrm{H}}$. This approach makes (25) difficult to solve, because the objective-function penalties that must be placed on the slack variables create numerical-scaling problems (Yagi, 2020). As such, we add the constraints:

$$w_{y,c,|\mathcal{T}^{\mathrm{H}}|}^{\mathcal{H},\mathrm{g}} + \min_{h \in \mathcal{T}^{\mathrm{H}}} \left\{ \sum_{h' \in \mathcal{T}^{\mathrm{H}}: h' \leq h} \left[\Upsilon_{y+1,h'} \cdot \left(W_{y+1,c,h'}^{\mathcal{H},f} - W_c^{\mathcal{H},-} \right) \right. \right.$$

$$+ \sum_{c' \in \bar{\mathcal{H}}_c} \Upsilon_{y+1,h'} W^{\mathcal{H},f}_{y+1,c',h'} \Bigg] \Bigg\} \geq 0; \forall c \in \mathcal{H}; \quad (29)$$

to (25), $\forall y \in \mathcal{Y}, y \neq |\mathcal{Y}|$. Constraint set (29) is an engineering estimate of the minimum water level that must be in each dam as of the end of each year (Conejo et al, 2006). The estimate that is in (29) is based on natural water inflows to each dam and its upstream dams and minimum water outflows from the dam. Our results show that (29) yields good computational performance and model fidelity, insomuch as our implementation of nested Benders's decomposition yields high-quality solutions compared to solving (1)–(23) directly.

To detail the algorithm that we use to implement our decomposition, $\forall y \in \mathcal{Y}$ we define Ω_y as the ordered set of optimality cuts that have been separated for the year-y variant of (25). Then, $\forall y \in \mathcal{Y}$, we can define the year-y subproblem with feasibility constraints and optimality cuts as:

$$\begin{aligned} \min_{a_y, \theta_y} \ & \Psi_y(a_y) + \theta_y \\ \text{s.t. } & a_y \in A_y(s_y) \\ & \theta_y \geq 0 \\ & (29) \\ & \theta_y \geq \Pi_{y+1}(\lambda^{\omega}_{y+1}, \eta(a_y)); \forall \omega \in \Omega_y; \end{aligned}$$

where λ^{ω}_{y+1} denotes the optimal dual-variable vector that is used to separate the ωth optimality cut in Ω_y. Hereafter, we denote this problem as $\mathcal{P}_y$.

Algorithm 1 provides pseudocode of our decomposition method. Figure 1 provides an accompanying visual depiction and shows (1)–(23) being decomposed into the sequence of year-1 through year-$|\mathcal{Y}|$ subproblems. The algorithm takes the following three inputs in Line 1: $\bar{\alpha}$ is the limit on the number of Benders's iterations to conduct, ϵ is the minimum desired optimality gap, and ζ is the threshold for dual-variable values that are used in separating optimality cuts. Line 2 initializes the algorithm. α, which is the iteration counter, is set equal to one. Ψ^{LB} and Ψ^{UB} are lower and upper bounds, respectively, on the optimal value of (1), and are initialized to $-\infty$ and $+\infty$, respectively. Finally, the optimality-cut sets, Ω_y, are initialized to be empty, $\forall y \in \mathcal{Y}$.

The main Benders's iterations are in Lines 3–18. Each iteration consists of two recursions. First, Lines 4–7 are the so-called forward simulation, which iterate through each $y \in \mathcal{Y}$ in time sequence, solve the year-y subproblem (*cf.* Line 5), and compute the associated year-y objective-function value (*cf.* Line 6). The forward recursion is depicted in Figure 1 by the 'forward' arrows that convey, $\forall y \in \mathcal{Y}, y \neq 1$, the starting state variable, s_y, of the year-y subproblem based on the solution that is obtained from solving the year-$(y-1)$ subproblem. Following the forward simulation, Line 8 updates the upper bound on the optimal value of (1). The value:

$$\sum_{y \in \mathcal{Y}} \Psi^{\alpha}_y;$$

Algorithm 1 Nested Benders's Decomposition

1: **input:** $\bar{\alpha}$, ϵ, ζ
2: **initialize:** $\alpha \leftarrow 1$, $\Psi^{\mathrm{LB}} \leftarrow -\infty$, $\Psi^{\mathrm{UB}} \leftarrow +\infty$, $\Omega_y \leftarrow \emptyset$, $\forall y \in \mathcal{Y}$
3: **repeat**
4: **for** $y \leftarrow 1$ **to** $|\mathcal{Y}|$ **do**
5: $(a_y^\alpha, \theta_y^\alpha) \leftarrow \arg\min \mathcal{P}_y$
6: $\Psi_y^\alpha \leftarrow \Psi_y(a_y^\alpha)$
7: **end for**
8: $\Psi^{\mathrm{UB}} \leftarrow \sum_{y \in \mathcal{Y}} \Psi_y^\alpha$
9: **for** $y \leftarrow |\mathcal{Y}|$ **to** 2 **do**
10: **if** $\theta_{y-1}^\alpha < \Psi_y^\alpha + \theta_y^\alpha$ **then**
11: $\tilde{\lambda}_y^\alpha \leftarrow \arg\max$ (27)
12: $\Omega_{y-1} \leftarrow \Omega_{y-1} \cup (|\Omega_{y-1}| + 1)$
13: $\lambda_y^{|\Omega_{y-1}|} \leftarrow \mathcal{I}(\tilde{\lambda}_y^\alpha, \zeta)$
14: **end if**
15: **end for**
16: $\Psi^{\mathrm{LB}} \leftarrow \Psi_1^\alpha + \theta_1^\alpha$
17: $\alpha \leftarrow \alpha + 1$
18: **until** $\alpha > \bar{\alpha}$ **or** $(\Psi^{\mathrm{UB}} - \Psi^{\mathrm{LB}})/\Psi^{\mathrm{LB}} \leq \epsilon$

gives a valid upper bound on (1) because the values of a_y^α, $\forall y \in \mathcal{Y}$ that are determined in Line 5 are feasible but not optimal in (1)–(23). A 'tighter' and monotonically non-increasing upper bound is given by:

$$\Psi^{\mathrm{UB}} \leftarrow \min \left\{ \Psi^{\mathrm{UB}}, \sum_{y \in \mathcal{Y}} \Psi_y^\alpha \right\}.$$

We use the upper bound that is given in Line 8 to examine more closely the impact of each optimality cut that is added as the algorithm progresses.

Lines 9–15 are the backward recursion. For each $y \in \mathcal{Y}$ such that θ_{y-1}^α underestimates $\Psi_y^\alpha + \theta_y^\alpha$ (*cf.* Line 10) year-y linear dual (27) is solved in Line 11 to obtain a dual-optimal vector, $\tilde{\lambda}_y^\alpha$. Next, the cut set, Ω_{y-1} is expanded by one in Line 12 and the cut coefficients are set in Line 13. The function, $\mathcal{I}(\cdot, \cdot)$, which appears in Line 13 imposes the threshold, ζ, on the cut coefficients. $\mathcal{I}(\cdot, \cdot)$ fixes equal to zero any element of $\tilde{\lambda}_y^\alpha$ that is smaller than ζ in magnitude. The backward recursion is depicted in Figure 1 by the 'backward' arrows, which, $\forall y \in \mathcal{Y}, y \neq |\mathcal{Y}|$, provide newly separated optimality cuts from the year-$(y+1)$ subproblem to the year-y subproblem. Line 16 updates the lower bound on the optimal value of (1). $\Psi_1^\alpha + \theta_1^\alpha$ is valid lower bound on (1) because this value is obtained by solving the year-1 subproblem, which can underestimate but cannot overestimate the optimal value of (1).[1] By principle of

[1] In theory, it is possible for $\Psi_1^\alpha + \theta_1^\alpha$ to overestimate the optimal value of (1), because we use $\mathcal{I}(\cdot, \cdot)$ to fix equal to zero any dual-variable value that has a magnitude below a certain threshold. We do not observe such behavior in our case study, however, because we select a sufficiently small value of ζ.

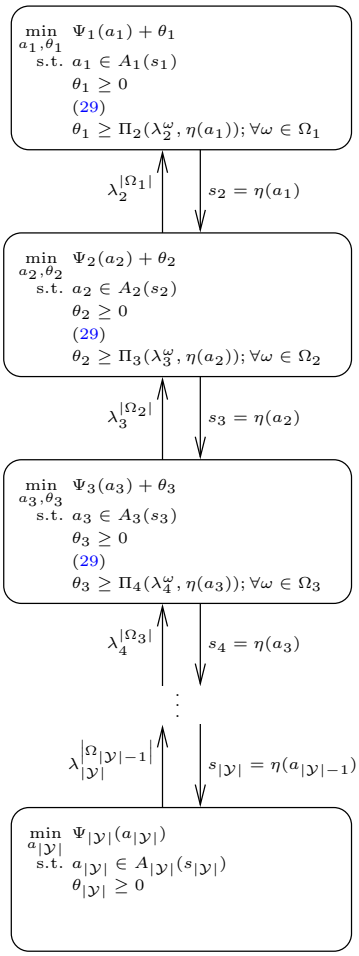

Fig. 1 Graphical illustration of decomposition technique and Algorithm 1.

optimality, the lower-bound that is computed in Line 16 is guaranteed to be monotonically non-decreasing. Line 17 updates the iteration counter. Line 18 terminates the algorithm once the iteration limit is exhausted or the optimality gap is below the threshold, ϵ.

We can guarantee finite convergence of Algorithm 1 to a global optimum by appealing to the work of Benders (1962). Specifically, $\forall y \in \mathcal{Y}$, the feasible region of (27) has a finite number of extreme points and extreme rays. Thus, there is a finite number of optimality cuts that can be added to $\mathcal{P}_y$, $\forall y \in \mathcal{Y}$. Once all of these cuts are added, solving $\mathcal{P}_y$, $\forall y \in \mathcal{Y}$ gives a global optimum of (1)–(23). In practice, Benders's decomposition finds a solution that is near a global optimum of (1)–(23) well before separating every possible optimality cut. Thus, if the termination criterion in Line 18 is used with $\bar{\alpha} < +\infty$ or $\epsilon > 0$, Algorithm 1 is likely to yield such a near-optimal solution.

4 Case-Study Data and Implementation

Our case study optimizes capacity expansion for Western Electricity Coordinating Council (WECC). Hydroelectric resources provide approximately a quarter of WECC's electric energy and generating capacity. The Columbia River system provides about half of WECC's hydroelectric capacity.[2] Our case study considers a 20-year optimization horizon that begins as of 2015. Most of the case-study data are gathered from WECC's 2024 and 2026 Production Cost Model Common Cases (PCMCCs).[3] Hourly load profiles are obtained from 2026 PCMCC, which provides actual and simulated load profiles for 2009 and 2026, respectively. We estimate a constant interannual load-growth factor from these data and apply the factor to interpolate and extrapolate the hourly load data to generate load profiles for each year of the case study. We assume a \$5 000/MWh penalty for load curtailment and a 7% annual discount rate (GE Energy, 2010). We apply 30% and 5% (on an energy basis) renewable-portfolio standards for wind and solar, respectively, to the final year of the optimization horizon.

4.1 Non-Hydroelectric-Generator Data

Our case study allows for investment in six generation technologies: coal-fired; natural-gas-fired steam (NGS), open-cycle (NGOC), and combined-cycle (NGCC); solar photovoltaic (PV); and wind units. We allow for retirement of these fossil-fueled technologies, but not of the renewable generation. In addition to these six technologies, we include hydroelectric, nuclear, geothermal, and biomass units, but assume that the capacities of these technologies are fixed. We use data from 2024 PCMCC to set starting capacity levels for the six generation technologies that allow for capacity additions and for the fixed capacities of the other technologies. We assume an 80% availability factor for nuclear, geothermal, and biomass units. We limit ramp rates to 0.3 p.u. for all technologies, except for hydroelectric units that are not in the Columbia River system, NGOC, PV, and wind, which are assumed to have ramp rates of 1.0 p.u. (Ibanez et al, 2014). We allow up to a 30% increase between two consecutive years in the installed capacity of any generation technology that allows for investment and the same maximum increase in the capacity of any transmission line. At most, 20% of the existing installed capacity of any technology that allows for retirement can be retired between two consecutive years. No energy storage, other than reservoirs that are in the Columbia River system, is considered.

Table 1 summarizes baseline generator-cost data (Black & Veatch Holding Company, 2012; Energy and Environmental Economics, Inc., 2012). We apply learning rates and regional multipliers to the values that are listed in the table, which allow for costs to vary by location and over time. Baseline capital costs of PV and wind decrease to \$2 226/kW and \$1 804/kW, respectively, by the end of the optimization horizon. Capital and fixed operation and maintenance costs of a given generation technology at individual electricity-system nodes are up to 20% higher or 15% lower than baseline costs. Generator-retirement costs are 5% of capital cost. Fuel costs of generators are computed using monthly estimates of coal and natural-gas prices for different WECC

[2] *cf.* https://www.wecc.org/Reliability/2015%20SOTI%20Final.pdf.
[3] *cf.* https://www.wecc.org/SystemAdequacyPlanning/Pages/Datasets.aspx.

regions, which are taken from 2024 PCMCC, and the heat rates that are reported in Table 1. We assume that a \$58/t carbon tax is levied on carbon emissions from combusting fossil fuels to generate electricity. Hourly available wind and solar generation is obtained from 2026 PCMCC. These wind- and solar-availability data are for a single year only. Thus, we apply the same availability profiles for all case-study years. Wind and solar are assumed to have zero marginal operating costs.

Table 1 Baseline generator-cost data for case study

Generation Technology	Capital Cost (\$/kW)	Fixed Operation and Maintenance Costs (\$/kW-year)	Variable Operation and Maintenance Costs (\$/MWh)	Heat Rate (Btu/kWh)
Coal	3 600	30.0	3.71	9 000
NGS	1 200	12.0	2.99	9 000
NGOC	1 150	12.0	2.99	9 200
NGCC	1 200	10.0	3.67	6 700
PV	3 325	50.0	0.00	n/a
Wind	2 000	60.0	0.00	n/a

4.2 Transmission Data

Transmission-network data are obtained from WECC's 2034 Reference Case. Figure 2 shows the topology of the 15-node transmission network. The states of Washington and Oregon are aggregated into the node that is labeled 'Pacific Northwest'. 2034 Reference Case includes different transmission limits that depend upon season and the direction of flow. We use the maximum capacity that is reported in the dataset for each line. We assume that adding transmission capacity incurs a cost of \$614 000/MW (Mason and Curry, 2012).

4.3 Hydroelectric Data

Figure 3 shows the topology of a 35-dam model of the Columbia River system. The 35-dam model is aggregated minimally (a dam name that is followed by '$+x$' indicates that the dam aggregates x additional dams). System-topology data are obtained from Northwest Power and Conservation Council[4] and Bonneville Power Administration. All of the water in the river system flows to Bonneville, from which it flows into Pacific Ocean. Circles in the figure indicate hydroelectric generators, squares indicate significant reservoirs. Dams without squares are run-of-river reservoirs that we assume have at most 3.5 hours of storage. Crosses indicate dams with non-trivial natural water inflows (*i.e.*, in addition to possible water inflows from an upstream dam). The green dotted rounded rectangles indicate the boundaries of a 19-dam aggregation that is used in most of our analysis. Results using the 19- and 35-dam models are qualitatively similar (Yagi, 2020; Yagi and Sioshansi, 2023). Dashed blue lines indicate state and

[4] *cf.* https://www.nwcouncil.org/energy/energy-topics/power-supply/.

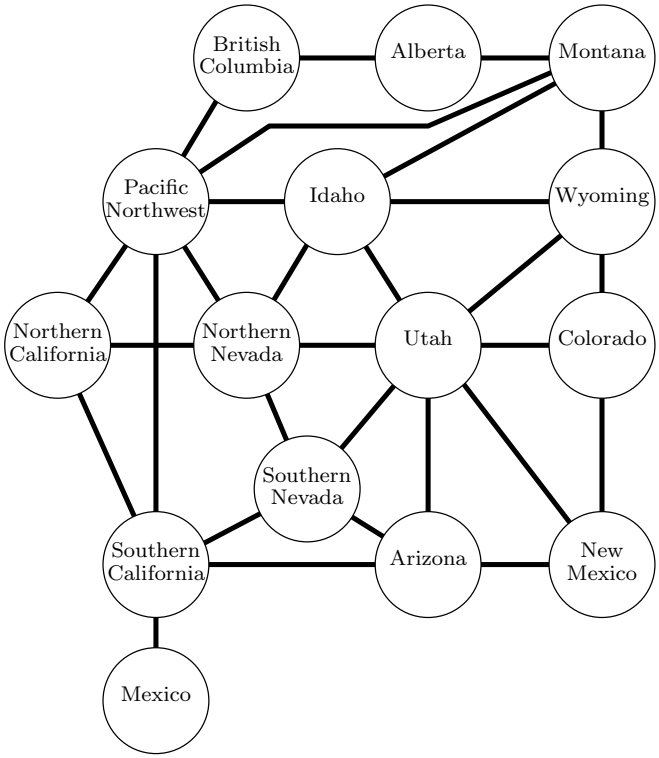

Fig. 2 15-node transmission-network topology for the case study.

provincial boundaries, and give a mapping between dams in the hydroelectric system to transmission-network nodes (*cf.* Figure 2).

Historical water-inflow, -outflow, and -storage and power-output data for each dam in the Columbia River system are obtained from United States Army Corps of Engineers[5] and Government of Canada.[6] Lower and upper bounds on water flows and reservoir levels are set to the minimum and maximum values that are observed in the historical data between 1999 and 2018. Historical water-inflow data for the same years are applied to the 20 years within the optimization horizon of the case study. The water-inflow data exhibit seasonal and interannual variability, which is expected.

All of the hydroelectric units that are located at a transmission node that are not part of the Columbia River system are aggregated into a single generic hydroelectric plant. Such plants are assumed to have fixed generation profiles, which are based on the historical generation profile of Lower Granite and scaled based on the aggregated nameplate capacity of the units. Lower Granite does not have a large upstream reservoir, meaning that its output reflects seasonal and annual variations of natural water inflows.

[5] *cf.* http://www.nwd-wc.usace.army.mil/dd/common/dataquery/www/.
[6] *cf.* https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html.

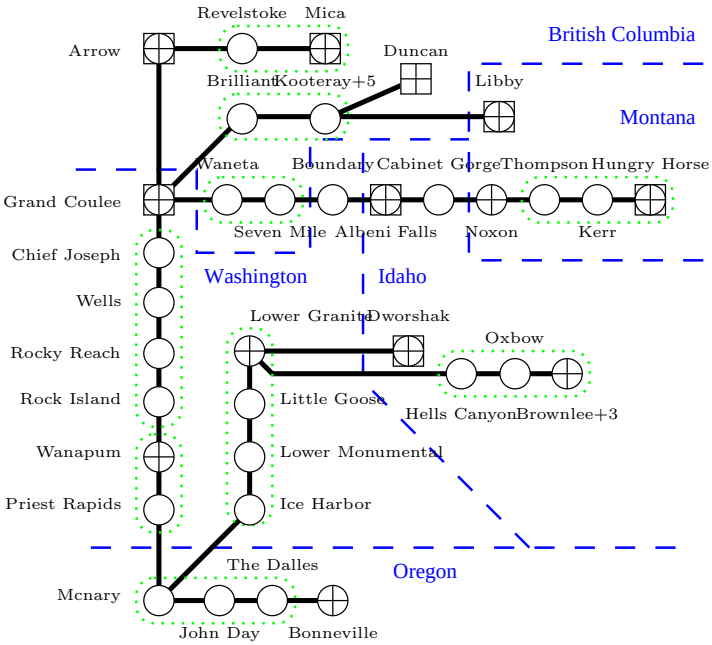

Fig. 3 Topology of Columbia River. Circles indicate hydroelectric generators, squares indicate significant reservoirs, crosses indicate dams with natural water inflows, and all water flows towards Bonneville. Green dotted rounded rectangles indicate boundaries of 19-dam aggregation of the system. Dashed blue lines indicate state and provincial boundaries.

4.4 Selection of Representative Operating Days

We populate the set, $\mathcal{T}^{\mathrm{H}}$, of hours that are used to represent system operations during each year by applying hierarchical clustering to the full data set for each year (Yagi and Sioshansi, 2023). To apply hierarchical clustering, we must define metrics between days and between clusters. To define such metrics, we represent each day as a vector, each of which consists of $24 \cdot |\mathcal{N}|$ load, solar-availability, and wind-availability features, up to $24 \cdot |\mathcal{H}|$ natural-water-inflow features for each dam with non-trivial natural water inflows that is part of the Columbia River system, and 24 generation features for the hydroelectric plants that are not part of Columbia River. Thus, the vector that describes each day has up to $24 \cdot (1 + 3 \cdot |\mathcal{N}| + |\mathcal{H}|)$ dimensions.

With this representation of each day, we use Euclidean distance and minimax linkage as the metrics between days and clusters, respectively (Bien and Tibshirani, 2011). Before clustering the data, all of the features are scaled linearly to the unit interval (*i.e.*, minimum and maximum values for each feature are re-scaled to 0 and 1, respectively). Minimax linkage provides each cluster's prototype, which is the day (from the unclustered data) that is closest to the cluster center. All of the days of the year that are within a single cluster are represented in (1)–(23) using the data that correspond to that cluster's prototype. We set the weight for each cluster equal to the number of days in the cluster. Using cluster prototypes (as opposed to cluster centroids) tends to provide better performance in decision making (Nahmmacher et al,

2016). Maintaining chronology of the cluster prototypes allows capturing seasonal variability in natural water inflows, load, and solar and wind availability.

4.5 Case-Study Implementation

Our case study is implemented on a system with 270 GB of memory and two Intel Xeon E5-2697 v4 processors, each of which has 18 2.30-GHz cores. All of the optimization models and Algorithm 1 are programmed with Python 2.7. The optimization models are solved using the barrier-method algorithm in Gurobi 7.5.1 with crossover disabled. The optimization models are too large to solve using Simplex method. Gurobi's presolver is disabled or set to conservative and the homogeneous-barrier-method algorithm is used and presolver aggregation is disabled in some cases to improve algorithm performance. The default optimality tolerance of 10^{-8} is used. However, in some cases Gurobi stops barrier iterations close to but before reaching this tolerance.

4.6 Assessing Solution and Algorithm Quality

We use three primary metrics to assess the quality of solutions that are obtained from our implementation of nested Benders's decomposition. The first is to examine the upper and lower bounds that are obtained in Lines 8 and 16 of Algorithm 1. These bounds, along with times that are required to apply Algorithm 1, measure the efficiency of the algorithm and the rate at which it obtains a near-optimal solution to (1)–(23).

Second, we use economic regret as a measure of the quality of the solutions that are found by Algorithm 1. Economic regret is computed using a three-step process (Ramírez-Sagner and Muñoz, 2019). First, (1)–(23) is solved using Algorithm 1. Next, the investment decisions that are obtained from applying Algorithm 1 are fixed and (1)–(23) is re-solved without applying Algorithm 1. This step gives us the total cost of making planning and operational decisions, if planning decisions are fixed based on Algorithm 1. Finally, this cost from solving (1)–(23) with the fixed investments is compared to the objective-function value from solving (1)–(23) with neither applying Algorithm 1 nor fixed investments. The difference between the objective-function values that are obtained in the second and third steps measures the cost (regret) of using Algorithm 1 to determine investment decisions as opposed to solving the undecomposed problem. Computing economic regret requires solving (1)–(23) without using Algorithm 1, which may be computationally intractable. In such cases, the objective-function value that is obtained from the second step of the process can be compared to the objective-function estimates that are obtained from applying Algorithm 1. In doing so, one can assess the extent to which Algorithm 1 provides a reasonable estimate of the 'true' objective-function value.

The final metric that we use is a direct comparison of the capacity levels that are determined from solving (1)–(23) with or without the application of Algorithm 1.

5 Case-Study Results

We begin by applying Algorithm 1 to an instance of (1)–(23) wherein hierarchical clustering is used to select ten representative operating days for each year. This problem instance can be solved directly without the use of Algorithm 1, which allows us to benchmark our proposed solution methodology to solving the undecomposed model directly. We use the crossover feature in the barrier algorithm that is used to solve the subproblems in Algorithm 1. Crossover applies Simplex method to the final solution that is obtained from the barrier algorithm to attempt to obtain a basic feasible solution to the problem. In doing so, crossover should provide better estimates of 'true' (primal- and) dual-variable values.

Figure 4 summarizes the convergence rate of Algorithm 1, by showing the upper and lower bounds that are obtained during each iteration and the corresponding optimality gap, as defined by Line 18 of the algorithm. The figure shows that the optimality gap closes quickly to about 1% within the first approximately 30 iterations. Thereafter, it takes considerable time to close the gap further. Algorithm 1 terminates after 200 iterations, which is the limit that we impose. Each iteration of Algorithm 1 requires an average of 10.8 minutes of wall-clock time. This can be contrasted with about 16 minutes of wall-clock time to solve the undecomposed problem. Algorithm 1 requires considerably more time to solve (1)–(23), compared to solving it directly. This is expected, given the relatively small problem instance.

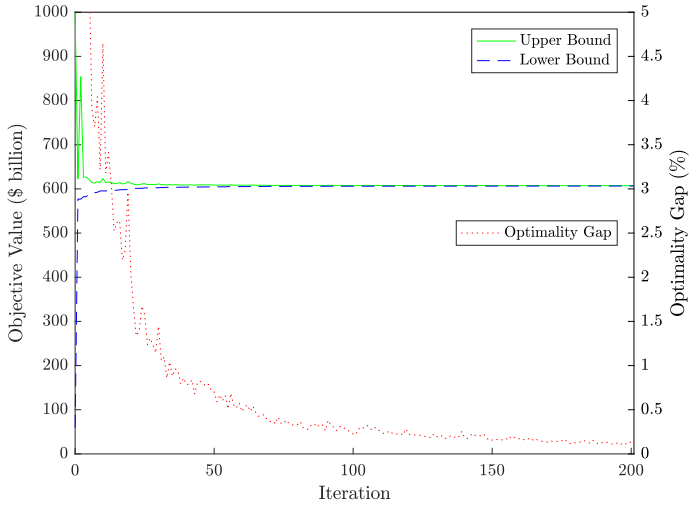

Fig. 4 Upper and lower bounds on optimal objective-function value and optimality gap of solutions that are obtained from applying Algorithm 1 with crossover to an instance of (1)–(23) with ten clusters to represent system operations.

Figure 5 summarizes the installed capacities of natural-gas-fired generation as of the final case-study year for solutions with different optimality gaps and for the solution that is obtained from solving (1)–(23) without decomposition. Figure 6 provides the same information regarding the installed capacities of transmission connectors.

Because there are many technologies and electricity-system nodes, we do not provide legends in the two figures. The purpose of the figures is to give a visual sense of how capacity decisions vary among different near-optimal solutions. Comparing the two figures shows that generator capacities are more varied than transmission capacities among the near-optimal solutions. These differences are particularly large for solutions with larger optimality gaps.

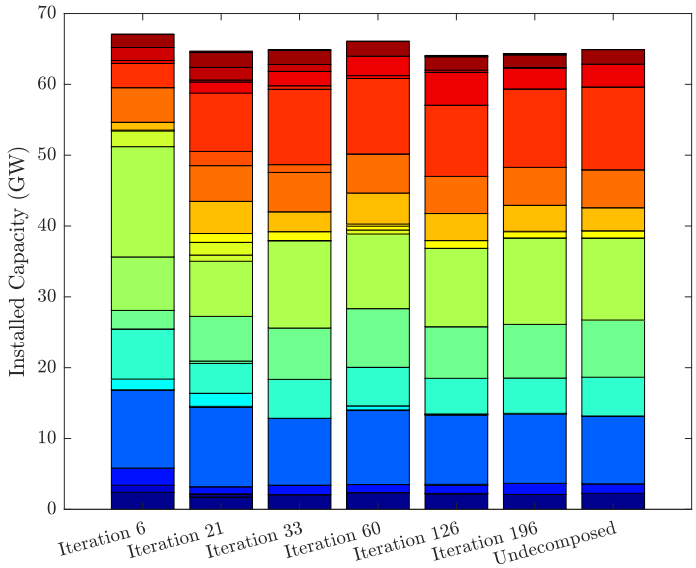

Fig. 5 Installed capacity of natural-gas-fired generation at different electricity-system nodes as of final case-study year of solutions that are obtained from applying Algorithm 1 with crossover to an instance of (1)–(23) with ten clusters to represent system operations.

Figure 7 summarizes economic regrets for the solutions that are summarized in Figs. 4–6. Economic regrets tend to be smaller than the optimality gaps that are reported in Fig. 4. For instance, the solution that is obtained from applying six iterations of Algorithm 1 has an optimality gap of 3.9% but an economic regret of appropriately 1%. These differences between the optimality gaps and economic regrets stem from two properties of the metrics. First, the upper and lower bounds that are shown in Fig. 4 represent the worst-case bounds on the optimal objective-function value. Thus, a solution with a particular optimality gap may be closer to optimal than the bounds suggest. Second, by its nature, (1)–(23) likely has many near-optimal solutions which may be very different in terms of capacity levels but quite similar in terms of cost. Indeed, Fig. 5 shows that applying six iterations of Algorithm 1 yields capacity decisions that are noticeably different from the solution that is obtained from solving the undecomposed problem. However, Fig. 7 shows that the six-iteration solution yields very small cost differences compared to a true optimum.

Using ten clusters likely provides a poor representation of system operations over the course of a full year (Yagi and Sioshansi, 2023). An issue with increasing the

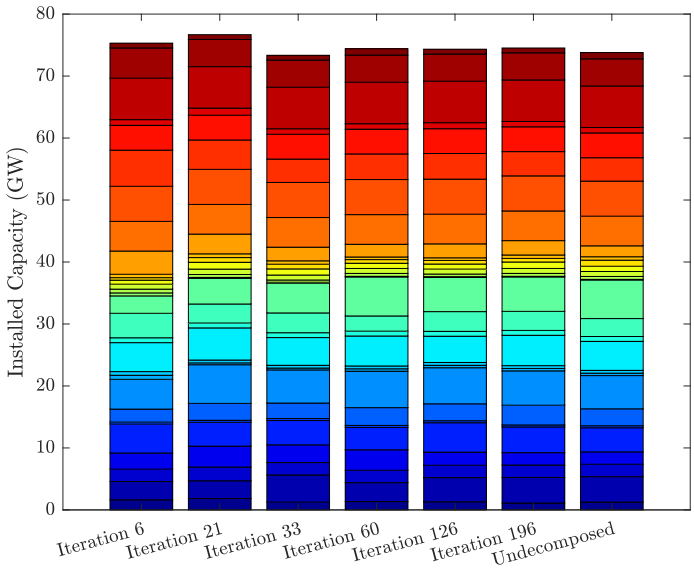

Fig. 6 Installed capacity of different transmission connectors as of final case-study year of solutions that are obtained from applying Algorithm 1 with crossover to an instance of (1)–(23) with ten clusters to represent system operations.

number of representative days is that the subproblems in Algorithm 1 grow in size and become more difficult to solve. In particular, crossover becomes more time-consuming. Figure 8 summarizes the convergence properties of applying Algorithm 1 without crossover to the same instance of (1)–(23) that is examined in Figs. 4–7.

Without crossover, the barrier method does not apply Simplex method to find a basic feasible solution to each subproblem. As such, each iteration of Algorithm 1 without crossover requires an average of 7.5 minutes of wall-clock time. Comparing Figs. 4 and 8 shows mixed relative algorithmic performance with and without crossover. More iterations (over 40 as opposed to 30) are required without crossover to achieve an optimality gap of 1% or less. However, Algorithm 1 without crossover finds a solution with an optimality gap of 0.1% or less (which is the termination criterion that we set) within 88 iterations. Using crossover provides more accurate dual-variable values, which yields more accurate optimality cuts. The less accurate optimality cuts that are obtained without crossover provide faster algorithmic convergence, because they tend to overestimate the values of θ_y, $\forall y \in \mathcal{Y}$.

Although Algorithm 1 converges more quickly without crossover as opposed to with, the less accurate optimality cuts may yield a lower-quality solution. Figure 9 summarizes the economic regrets of solutions that are obtained from applying Algorithm 1 without crossover to the same 10-day instance of (1)–(23). Contrasting Figs. 7 and 9 shows that the economic regrets of the solutions that are obtained without crossover are slightly higher than those that are obtained with crossover.

We conclude analysis of our case study by applying Algorithm 1 without crossover to an instance of (1)–(23) that includes all hours of every year. Figure 10 summarizes the convergence properties of applying Algorithm 1 without crossover to this problem

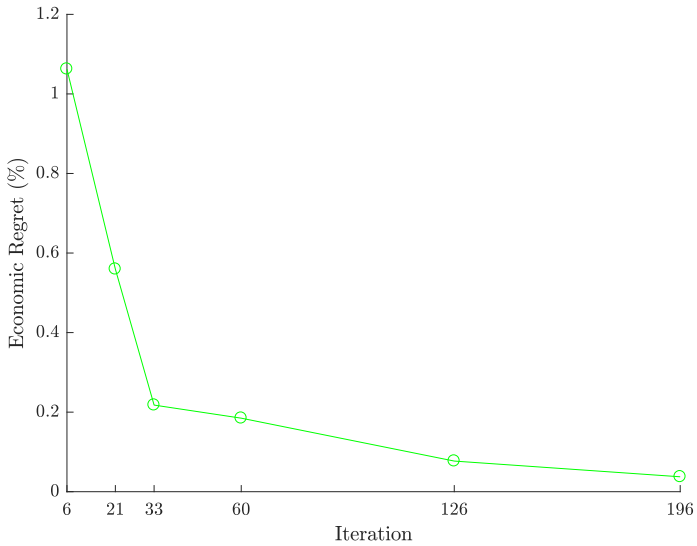

Fig. 7 Economic regrets of solutions that are obtained from applying Algorithm 1 with crossover to an instance of (1)–(23) with ten clusters to represent system operations.

instance. Each iteration of Algorithm 1 requires an average of 499.0 minutes and the algorithm obtains a solution with an optimality gap of 0.1% or less within 30 iterations. The undecomposed version of an instance (1)–(23) that includes all hours of every year cannot be solved, which demonstrates benefit of applying Algorithm 1 to a large problem instance.

Figures 11 and 12 summarize generation and transmission capacity levels as of the final case-study year for solutions that are obtained from applying Algorithm 1. As with Figs. 5 and 6, Figs. 11 and 12 show that the capacity levels converge and become more stable as more iterations of Algorithm 1 are performed to find solutions that are closer to optimal. Because the instance of (1)–(23) that includes every hour of each year cannot be solved without the application of Algorithm 1, we are unable to compute economic regrets.

6 Conclusions and discussions

This paper examines the challenging problem of capacity planning in electricity systems with significant penetrations of renewable and hydroelectric resources. Such modeling is difficult, because of the need to ensure that the electricity system has sufficient flexibility to maintain real-time balance between supply and demand. Hydroelectricity can provide such flexibility, because of its technical characteristics and the storage potential of reservoirs. However, modeling reservoirs requires capturing the chronology of decisions and seasonal and interannual variability of natural water inflows.

Our companion paper explores the use of three model reductions—simplifying hydroelectric-plant, hydroelectric-system, and operating-period representation—to

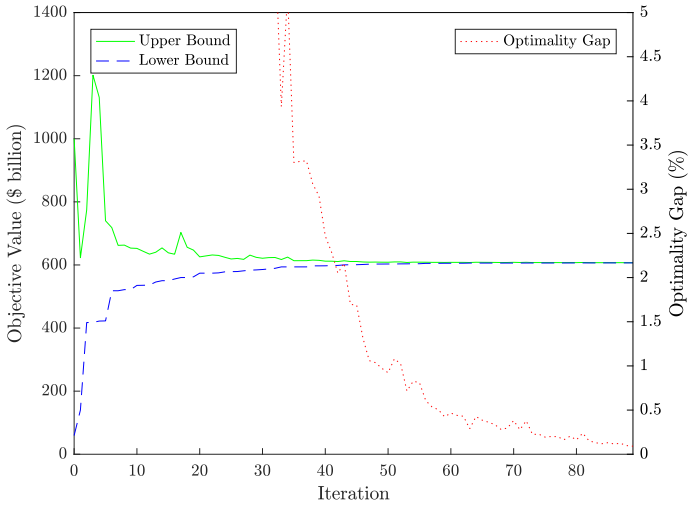

Fig. 8 Upper and lower bounds on optimal objective-function value and optimality gap of solutions that are obtained from applying Algorithm 1 without crossover to an instance of (1)–(23) with ten clusters to represent system operations.

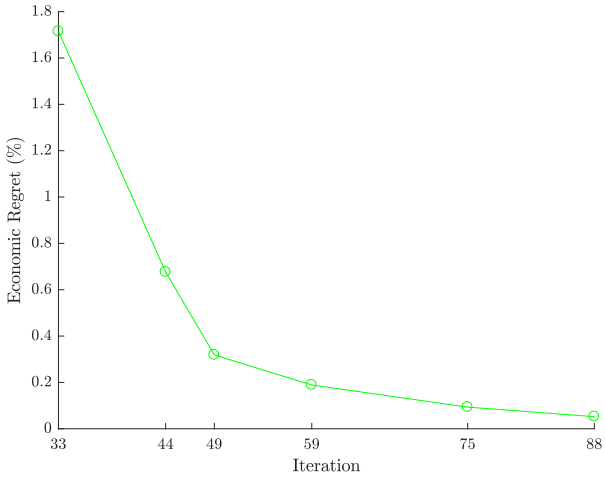

Fig. 9 Economic regrets of solutions that are obtained from applying Algorithm 1 without crossover to an instance of (1)–(23) with ten clusters to represent system operations.

improve model tractability (Yagi and Sioshansi, 2023). This paper explores the use of nested Benders's decomposition. The two works are complementary, insomuch as real-world capacity-planning problems likely will require the use of both types of techniques. Using economic regret, convergence rates, and capacity decisions as metrics, we assess the quality of investment decisions that are made using our implementation of the decomposition method. Our case study, which is based on WECC and

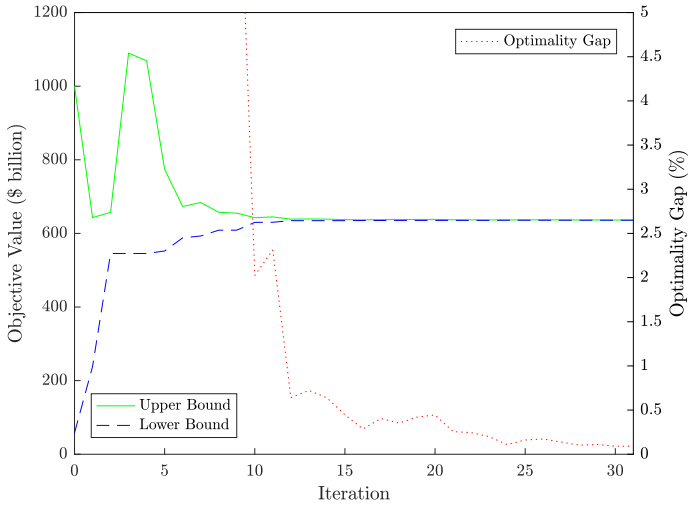

Fig. 10 Upper and lower bounds on optimal objective-function value and optimality gap of solutions that are obtained from applying Algorithm 1 without crossover to an instance of (1)–(23) with all hours of every year to represent system operations.

the Columbia River system, shows that the algorithm performs quite well. Indeed, we are able to solve to within less than a 0.1% optimality gap an instance of (1)–(23) with a 20-year planning horizon and representing every hour of every year. Without decomposition, such a problem instance exhausts the memory of the 270-GB system on which our computations are conducted.

Our investment model is deterministic and assumes that all decisions are continuous. Nested Benders's decomposition can be extended naturally to include stochasticity, using what is known as stochastic dual dynamic programming (SDDP) (Pereira and Pinto, 1991). Important sources of stochasticity in our model can include uncertain hydrological conditions (*e.g.*, dry versus wet years) and policy or technology changes or developments. The key distinction between Algorithm 1 and SDDP is that the random variables are sampled within the SDDP algorithm and used in the forward simulation and backward recursion. Including discrete variables in our model is a greater challenge as opposed to stochasticity. The challenge arises from the fact that discrete variables eliminate the strong-duality property on which optimality cuts are based and raise challenges with finding dual variables. Extending Benders's decomposition to problems with discrete variables is an area of active research (Zou et al, 2019).

Acknowledgments. The authors thank the editors, two anonymous reviewers, A. J. Conejo, R. C. Perez, and A. Sorooshian for helpful discussions, comments, and suggestions. Any errors, opinions, and conclusions that are expressed in this paper are solely those of the authors.

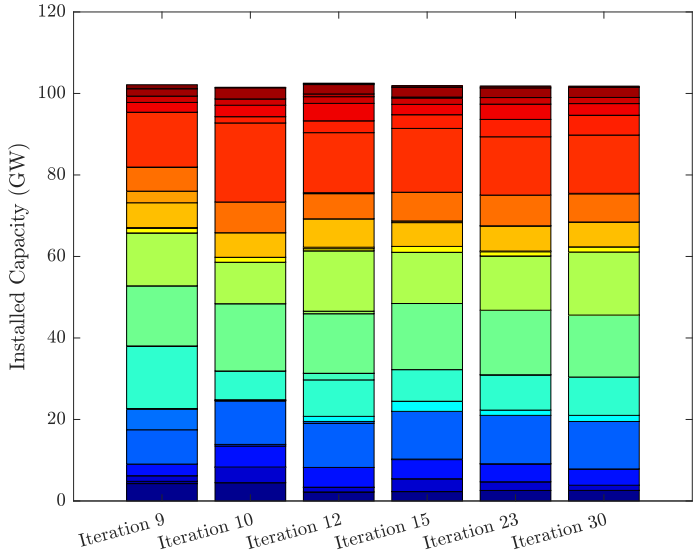

Fig. 11 Installed capacity of natural-gas-fired generation at different electricity-system nodes as of final case-study year of solutions that are obtained from applying Algorithm 1 without crossover to an instance of (1)–(23) with all hours of every year to represent system operations.

Declarations

- Funding: Open access funding provided by Carnegie Mellon University. This work was supported by Tokushu Tokai Paper Co., Ltd., which helped to fund the first author's Ph.D. studies at The Ohio State University, National Science Foundation grants 1463492, 1808169, and 1922666, and Carnegie Mellon Electricity Industry Center.

References

Ahlhaus P, Stursberg P (2013) Transmission Capacity Expansion: An Improved Transport Model. In: 2013 4th IEEE/PES Innovative Smart Grid Technologies Europe (ISGT EUROPE), Institute of Electrical and Electronics Engineers, Lyngby, Denmark

Barrera-Santana J, Sioshansi R (2023) An Optimization Framework for Capacity Planning of Island Electricity Systems. Renewable and Sustainable Energy Reviews 171:112955

Benders JF (1962) Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik 4:238–252

Bertsekas DP (1995) Nonlinear Programming, 2nd edn. optimization and computation, Athena Scientific, Belmont, Massachusetts

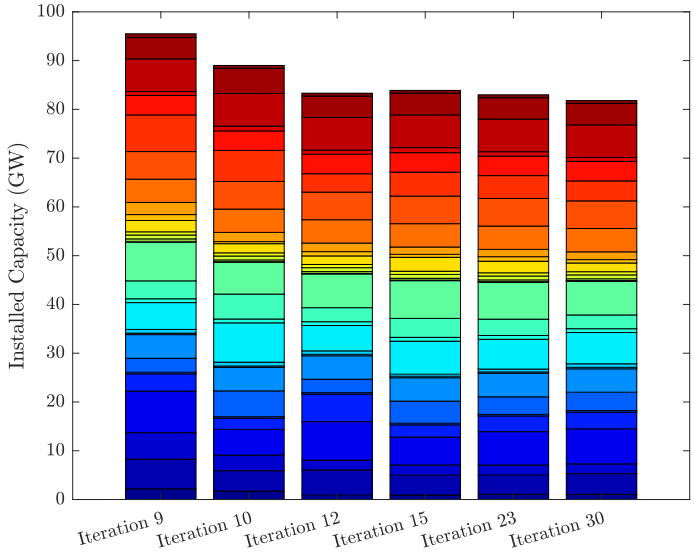

Fig. 12 Installed capacity of different transmission connectors as of final case-study year of solutions that are obtained from applying Algorithm 1 without crossover to an instance of (1)–(23) with all hours of every year to represent system operations.

Bien J, Tibshirani R (2011) Hierarchical Clustering With Prototypes via Minimax Linkage. Journal of the American Statistical Association 106:1075–1084

Birge JR (1985) Decomposition and Partitioning Methods for Multistage Stochastic Linear Programs. Operations Research 33:989–1007

Black & Veatch Holding Company (2012) Cost and Performance Data for Power Generation Technologies. Black & Veatch Holding Company, prepared for National Renewable Energy Laboratory

Boffino L, Conejo AJ, Sioshansi R, et al (2019) A Two-Stage Stochastic Optimization Planning Framework to Deeply Decarbonize Electric Power Systems. Energy Economics 84:104457

Cohen S, Becker J, Bielen D, et al (2019) Regional Energy Deployment System (ReEDS) Model Documentation: Version 2018. Tech. Rep. NREL/TP-6A20-72023, National Renewable Energy Laboratory, Golden, CO

Conejo AJ, Castillo E, Mínguez R, et al (2006) Decomposition Techniques in Mathematical Programming, 1st edn. Springer-Verlag Berlin, Heidelberg, Germany

Energy and Environmental Economics, Inc. (2012) Cost and Performance Review of Generation Technologies: Recommendations for WECC 10-and 20-Year Study Process. Energy and Environmental Economics, Inc., prepared for Western Electric Coordinating Council

Falugi P, Konstantelos I, Strbac G (2016) Application of Novel Nested Decomposition Techniques to Long-Term Planning Problems. In: 2016 Power Systems Computation Conference, Institute of Electrical and Electronics Engineers, Genoa, Italy

GE Energy (2010) Western Wind and Solar Integration Study. Tech. Rep. NREL/SR-550-47434, National Renewable Energy Laboratory, Golden, CO

Hidalgo IG, Fontane DG, Lopes JEG, et al (2014) Efficiency Curves for Hydroelectric Generating Units. Journal of Water Resources Planning and Management 140:86–91

Hjelmeland MN, Helseth A, Korpås M (2019) Medium-Term Hydropower Scheduling with Variable Head under Inflow, Energy and Reserve Capacity Price Uncertainty. Energies 12:189

Huertas-Hernando D, Farahmand H, Holttinen H, et al (2017) Hydro power flexibility for power systems with variable renewable energy sources: an IEA Task 25 collaboration. WIREs Energy and Environment 6:e220

Hunter-Rinderle RG, Sioshansi R (2021) Data-Driven Modeling of Operating Characteristics of Hydroelectric Generating Units. Current Sustainable/Renewable Energy Reports 8:199–206

Ibanez E, Magee T, Clement M, et al (2014) Enhancing hydropower modeling in variable generation integration studies. Energy 74:518–528

Lara CL (2019) Decomposition Algorithms for Optimal Manufacturing and Power Systems Infrastructure Planning. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania

Liu Y, Sioshansi R, Conejo AJ (2018a) Hierarchical Clustering to Find Representative Operating Periods for Capacity-Expansion Modeling. IEEE Transactions on Power Systems 33:3029–3039

Liu Y, Sioshansi R, Conejo AJ (2018b) Multistage Stochastic Investment Planning with Multiscale Representation of Uncertainties and Decisions. IEEE Transactions on Power Systems 33:781–791

Liu Y, Hunter-Rinderle RG, Luo C, et al (2021) How Climate-Related Policy Affects the Economics of Electricity Generation. Current Sustainable/Renewable Energy Reports 8:17–30

Luenberger DG, Ye Y (2008) Linear and Nonlinear Programming, International Series in Operations Research & Management Science, vol 116, 3rd edn. Springer, New York, New York

Maceiral MEP, Penna DDJ, Diniz AL, et al (2018) Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil-Main Features and Improvements on the NEWAVE Model. In: 2018 Power Systems Computation Conference, Institute of Electrical and Electronics Engineers, Dublin, Ireland

Maluenda B, Negrete-Pincetic M, Olivares DE, et al (2018) Expansion planning under uncertainty for hydrothermal systems with variable resources. International Journal of Electrical Power & Energy Systems 103:644–651

Mason T, Curry T (2012) Capital Costs for Transmission and Substations: Recommendations for WECC Transmission Expansion Planning. B&V Project Number 176322, prepared for Western Electricity Coordination Council

Merrick JH (2016) On representation of temporal variability in electricity capacity planning models. Energy Economics 59:261–274

Nahmmacher P, Schmid E, Hirth L, et al (2016) Carpe diem: A novel approach to select representative days for long-term power system modeling. Energy 112:430–442

Pereira MVF, Pinto LMVG (1985) Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach. Water Resources Research 21:779–792

Pereira MVF, Pinto LMVG (1991) Multi-stage stochastic optimization applied to energy planning. Mathematical Programming 52:359–375

Poncelet K, Höschle H, Delarue E, et al (2017) Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems. IEEE Transactions on Power Systems 32:1936–1948

Rahmaniani R, Crainic TG, Gendreau M, et al (2017) The Benders decomposition algorithm: A literature review. European Journal of Operational Research 259:801–817

Raineri R (2006) Chile: Where It All Started. In: Sioshansi FP, Pfaffenberger W (eds) Electricity Market Reform: An International Experience. Elsevier

Ramírez-Sagner G, Muñoz FD (2019) The effect of head-sensitive hydropower approximations on investments and operations in planning models for policy analysis. Renewable and Sustainable Energy Reviews 105:38–47

Sioshansi R (2016) Retail Electricity Tariff and Mechanism Design to Incentivize Distributed Renewable Generation. Energy Policy 95:498–508

Sioshansi R, Conejo AJ (2017) Optimization in Engineering: Models and Algorithms, Springer Optimization and Its Applications, vol 120. Springer Nature,

Gewerbestraße 11, 6330 Cham, Switzerland

Sioshansi R, Denholm P, Arteaga J, et al (2022) Energy-Storage Modeling: State-of-the-Art and Future Research Directions. IEEE Transactions on Power Systems 37:860–875

Stoft S (2002) Power System Economics: Designing Markets for Electricity. Wiley-Interscience, New York, New York

Yagi K (2020) Analyses of Issues Arising in Power Systems and Electricity Markets with High Renewable Penetration. PhD thesis, The Ohio State University, Columbus, Ohio

Yagi K, Sioshansi R (2021) Do Renewables Drive Coal-Fired Generation Out of Electricity Markets? Current Sustainable/Renewable Energy Reports 8:222–232

Yagi K, Sioshansi R (2023) How Market Power Can Suppress the Effect of Carbon Policies in Wholesale Electricity Markets. Journal of Japan Society of Energy and Resources 44:211–219

Zou J, Ahmed S, Sun XA (2019) Stochastic dual dynamic integer programming. Mathematical Programming 175:461–502