Noname manuscript No.
(will be inserted by the editor)

Electric Power Infrastructure Planning Under Uncertainty: Stochastic Dual Dynamic Integer Programming (SDDiP) and parallelization scheme

Cristiana L. Lara · John D. Siirola · Ignacio E. Grossmann

Received: date / Accepted: date

Abstract We address the long-term planning of electric power infrastructure under uncertainty. We propose a Multistage Stochastic Mixed-integer Programming formulation that optimizes the generation expansion to meet the projected electricity demand over multiple years while considering detailed operational constraints, intermittency of renewable generation, power flow between regions, storage options, and multiscale representation of uncertainty (strategic and operational). To be able to solve this large-scale model, which grows exponentially with the number of stages in the scenario tree, we decompose the problem using Stochastic Dual Dynamic Integer Programming (SDDiP). The SDDiP algorithm is computationally expensive but we take advantage of parallel processing to solve it more efficiently. The proposed formulation and algorithm are applied to a case study in the region managed by the Electric Reliability Council of Texas (ERCOT) for scenario trees considering natural gas price and carbon tax uncertainty for the reference case and a hypothetical case without nuclear power. We show that the parallelized SDDiP algorithm allows the solution of multistage stochastic programming models with quadrillions of variables and constraints in reasonable amounts of time.

1 Introduction

Changes in electricity demand, together with the wear-and-tear and retirement of old generators, and the advances in the technology pool for electricity generation and

Cristiana L. Lara · Ignacio E. Grossmann
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA, USA E-mail: cristianallara@cmu.edu

John D. Siirola
Center for Computing Research
Sandia National Laboratories
Albuquerque, NM, USA

storage, make it necessary to expand or adapt the electric power infrastructure. Generation expansion planning (GEP) models can be used to support these investment decisions, as well as to study the impact of new technology developments, resource cost trends, and policy shifts (e.g. carbon tax, minimum renewable generation quota) [77, 37, 2, 20].

Power systems are subject to a variety of systematic uncertainties such as fuel prices, load demand, renewable generation, disruptive technologies, and future policies. However, because of the computational expense of combining uncertainty with a complete representation of the grid, and integrating detailed operating decisions with investment decisions over long planning horizons [81, 62, 70, 61, 52], most of the available commercial tools [48, 49, 80, 15] and academic models [16, 80, 68, 36, 18, 28, 39] are deterministic. The body of literature that addresses GEP optimization problem under uncertainty can be classified into two fundamental approaches for capturing uncertainty: Robust Optimization and Stochastic Programming.

The main idea behind Robust Optimization (RO) is to guarantee feasibility over a specified uncertainty set by modeling uncertain variables using bounds [5]. In general, this means that the computational burden of RO is much lower than that of stochastic programming. However, RO predicts more conservative results compared to the latter [24]. Malcolm and Zenios [51] were the first to propose a RO model for power systems capacity planning. Since then, other authors have explored GEP formulations in the context of RO [58, 12, 43, 42]. More recently, Adjustable Robust Optimization (ARO) has risen to prominence as an alternative to classic RO. ARO introduces recourse into the traditional RO formulation, allowing the model to respond to some uncertainty and generate less conservative solutions [4, 87]. Some of the papers addressing GEP with ARO are [53, 54, 56, 3].

Stochastic Programming (SP) is the most popular modeling framework for GEP problems under uncertainty. SP ~~tends to be more appropriate for long-term production planning and strategic design decisions because it allows recourse decisions in the future to adapt to how the uncertainties are revealed [24]. Thus, it~~ is less conservative than RO as it typically limits the feasibility requirement to a finite number of realization instead of requiring feasibility for a continuous region of parameter (uncertainty set). SP assumes uncertain data are random variables with known probability distributions and uses sampled values from this distribution to build a scenario tree and optimize over the expectation [7]. As a downside, the solution is dependant on the accuracy of the assumed probability distributions of the uncertain parameters. However, there is mathematical theory and computational evidence that solutions obtained from SP are often stable with respect to changes in input probability distributions [73].

SP models can be formulated as two-stage and multistage problems. A typical two-stage stochastic GEP model considers as first-stage *here-and-now* decisions the investment decision over the entire planning horizon, which is made before the uncertainty realization. In this context, the second-stage *wait-and-see* decisions are the operational decisions, which are fully adaptive to the uncertainty realization. Some of the GEP literature that formulates the problem as two-stage stochastic programming includes [14, 1, 47, 33, 85, 60, 21, 32, 59, 22].

Multistage stochastic programming GEP models allow recourse between investment decisions in each stage, hence, they are also fully adaptive to the uncertainty realization. Park and Baldick [64] propose a multistage stochastic mixed-integer program to solve GEP under load and wind availability uncertainty, and solve the model using a rolling-horizon. Zhan et al. [86] propose a multistage stochastic programming model with endogenous uncertainty for GEPs with large amounts of wind power, and introduce a quasi-exact solution approach to reformulate the model as a mixed-integer linear programming model. Liu et al. [44] propose a multistage linear stochastic GEP model that captures both large-scale uncertainties (e.g., investment, fuel-cost demand-growth rate uncertainty), and small-scale uncertainties (e.g., hourly demand and renewable generation uncertainty), and use progressive hedging algorithm to decompose the model by scenario and reduce computation times. Zou et al. [88] propose a partially adaptive stochastic mixed-integer optimization model in which the capacity expansion is fully adaptive to uncertainty up to a certain period and follows a two-stage approach thereafter, and propose an approximation algorithm to solve their model efficiently.

As mentioned before, even deterministic GEP models can pose significant computational challenges as the temporal and spatial scale resolution are increased ~~[39]~~ [46,39]. The added complexity of handling uncertainty greatly intensifies this challenge, especially for multistage stochastic programming formulations as the scenario tree grows exponentially with the number of stages. Therefore, significant research has been devoted to the development of decomposition techniques to allow the solution of these problems in an efficient matter.

The most popular decomposition methods applied to multistage stochastic programming problems can be classified as scenario-based (e.g., Lagrangean Decomposition [25], Dual Decomposition [9,35], Augmented Lagrangean [57] and Progressive Hedging [84]), and stage-based decomposition (e.g., Nested Benders Decomposition [6], Stochastic Dual Dynamic Programming (SDDP) [66], Stochastic Dual Dynamic integer Programming (SDDiP) [89]). Both categories have guaranteed finite convergence for linear programming (LP) formulations. However, for the case of mixed-integer linear programming (MILP) formulations they can provide bounds to the optimal solution, but generally do not have guaranteed finite convergence.

Scenario-based decomposition often utilizes the framework of Lagrangean decomposition to decompose the problem into scenarios by dualizing the nonanticipativity constraints. Examples of GEP decompositions that fit within this category are [33,59,44]. Stage-base decomposition decomposes the model by nodes in the scenario tree, and are usually based on Benders decomposition. Stage-based algorithms have smaller subproblems and are more suitable for models in which the solution of a single scenario is already computationally very demanding. SDDP has been widely used in the context of optimal scheduling of hydrothermal generating systems [65,66, 76,82]. Regarding GEP problems, both Nested Decomposition and SDDP have been used in combination with Benders decomposition for two-stage stochastic programming models in which the operational subproblems are challenging [23,74]. SDDiP [89], which is an extension of SDDP to multistage integer programming models, is a promising technique to solve multistage stochastic integer programming models that has great potential for GEP problems with hourly operating details.

In this paper, we address the long-term planning of electric power infrastructure under multiscale uncertainty. We propose a Multistage Stochastic Mixed-Integer Programming (MSIP) formulation that takes the viewpoint of a central planning entity, whose goal is to optimize the generation expansion to meet the projected electricity demand over a long-planning horizon, while considering multiple sources (natural gas, coal, nuclear, solar, wind and storage), detailed operational constraints on an hourly basis, variability and intermittency of renewable generation sources, power flow between regions, and multi-scale uncertainty (i.e. investment level uncertainty - e.g. fuel price and carbon tax uncertainties - and operating level uncertainty - e.g. renewable availability and hourly load demand). The model is an extension of the deterministic model by Lara et al. [39] and can be of the order of quadrillion variables and constraints. To be able to solve such a large-scale model, we decompose the problem using Stochastic Dual Dynamic Integer Programming (SDDiP) [89] and take advantage of parallel processing to solve it more efficiently. The framework presented here is slightly different than the one proposed by [89] as it allows mixed-integer recourse at the expense of losing the finite convergence property (i.e. there is a potential dual gap). This is first discussed by Lara et al. [39], which presents three options of cuts for the deterministic version of the Nested Decomposition and prove their validity (not their tightness) in the context of mixed-integer state variables.

The proposed GEP model follows four out of the five trends listed by Babatunde et al.'s survey on GEP [2]: (i) it handles uncertainty; (ii) it considers renewable energy penetrations and includes short-term operating decisions; (iii) it includes the option of adding energy storage; (iv) it addresses sustainable issues by having the option of imposing minimum renewable generation quota, maximum CO_2 emissions quota, and/or carbon tax. The only trend from [2] that this paper does not address is the deregulation of the power sector.

The major contributions of this paper are the following: (i) application of SDDiP in the context of GEP optimization with integrated operating decisions, (ii) SDDiP with mixed-integer recourse, (iii) parallelization scheme to solve the SDDiP more efficiently, and (iv) application of the model and algorithmic framework to a case-study for Electric Reliability Council of Texas (ERCOT) region considering operational and strategic uncertainty.

The remainder of the paper is organized as follows: Section 2 presents the problem statement, discusses modeling assumptions, proposes a concise representation of the multistage mixed-integer linear programming model for GEP optimization under uncertainty (the detailed model is shown in Appendix A), and discusses how the scenario tree is generated. Section 3 describes the SDDiP algorithm, its framework, major assumptions, and how we parallelize the algorithm. In Section 4 we first present the results for a case-study for the ERCOT region, showing the order of magnitude of problems the SDDiP algorithm can solve on a personal computer and the solution time. Then, we compare the first-stage *here-and-now* decisions for the reference case with natural gas price uncertainty, the *no nuclear* case with natural gas price uncertainty, carbon tax uncertainty, and high carbon tax uncertainty, and show the value of stochastic programming for the *no nuclear* case with high carbon tax uncertainty. Finally, in Section 5 we draw some conclusions.

2 Formulation

The proposed GEP problem involves choosing the optimal investment strategy and operating schedule for the power system in order to meet the projected load demand over the time-horizon for each location, while minimizing the expected net present cost over the scenario tree. This is an extension of the MILP problem proposed by [39] to multistage stochastic mixed-integer programming, in order to address uncertainty.

A set of existing and potential generators is given, for which the energy source (nuclear, coal, natural gas, wind or solar)[1] and the generation technology are known.

- For the existing generators we consider: (a) **coal:** steam turbine (coal-st-old); (b) **natural gas:** boiler plants with steam turbine (ng-st-old), combustion turbine (ng-ct-old), and combined-cycle (ng-cc-old); (c) **nuclear:** steam turbine (nuc-st-old); (d) **solar:** photovoltaic (pv-old); (e) **wind:** wind turbine (wind-old);
- For the potential generators we consider: (a) **coal:** without (coal-new) and with carbon capture (coal-ccs-new); (b) **natural gas:** combustion turbine (ng-ct-new), combined-cycle without (ng-cc-new) and with carbon capture (ng-cc-ccs-new); (c) **nuclear:** steam turbine (nuc-st-new); (d) **solar:** photovoltaic (pv-new) and concentrated solar panel (csp-new); (e) **wind:** wind turbine (wind-new);

Also known are: their nameplate (maximum) capacity; lifetime; fixed and variable operating costs; start-up cost (fixed and variable); cost for extending their lifetimes; CO_2 emission factor and carbon tax, if applicable; fuel price; and operating characteristics such as ramp-up/ramp-down rates, operating limits, and contribution to spinning and quick start fraction for thermal generators.

For the case of existing generators, their age at the beginning of the time-horizon and location are also known. For the case of potential generators, the capital cost (which is a linear function of its nameplate capacity), and the maximum yearly installation of each generation technology are also given. Also given is a set of potential storage units, with specified technology (we consider as options lithium-ion, lead-acid, and flow batteries), capital cost, power rating, rated energy capacity, charge and discharge efficiency, and storage lifetime. Additionally, the multiple profiles of projected load demand and renewable availability (capacity factor) are given for each location, as well as the distance between locations, the transmission loss per mile, and the transmission line capacity between locations.

The problem is then to find the optimal "here-and-now" decisions for the first-stage and "wait-and-see" decisions for the remaining stages and respective scenarios regarding: a) location, year, type, and number of generators and storage units to install; b) when to retire generators and storage units; c) whether or not to extend the life of the generators that reach their expected lifetime; d) an approximate operating schedule for each installed generator; and e) the approximate power flow between each location in order to meet the projected demand. The goal is to minimize the expected net present cost over the scenario tree (including operating, investment, and

[1] In this paper we do not consider hydroelectric power as it is available in very limited amounts in the ERCOT region.

environmental costs). The large-scale strategic level uncertain parameters are user-defined and can be, for example, the yearly fuel price, and any potential carbon tax. The small-scale operating level uncertain parameters (renewable generation availability and load demand) are captured through multiple representative days between investment decisions.

2.1 Modeling assumptions

In order to improve tractability and allow the solution of the MSIP model for large areas over a few decades with multiple scenarios, we adopted judicious modeling aggregations and approximations to address the multi-scale aspects, both in its spatial and temporal dimensions [39]. In order to significantly reduce computation time, generator clustering [63] and time sampling [39] approaches are adopted.

The area considered is divided into regions that have similar climate (e.g., wind speed and solar incidence over time), and load demand profiles. It is assumed that the potential locations for the generators and storage units are the midpoints of each region r. Additionally, generators and storage units that have the same characteristics, such as technology and operating status (i.e., existing or potential), are aggregated into clusters i and j, respectively, for each region r [63]. The major impact of this approximation in the model formulation is that the discrete variables associated with generators and storage units correspond to integer rather than binary variables to represent the number of generators/storage units under a specific status in cluster i and j, respectively.

~~We use~~ As discussed extensively in the literature (e.g. [81,16,62,68,28,52]), a system with increasing renewable penetration has the need for chronological hourly or sub-hourly representation to better address the variability in renewable energy generation and better assess the trade-offs between long-term investment decisions and short-term operating decisions. We use representative days as a compromise between time resolution and model tractability. We use the same representative days as [39], selected from historical data via k-means clustering approach, where the goal of the clustering procedure is to select representative days to approximate: (i) the duration curves of historical load and renewables time series, (ii) the temporal correlation of each time series, and (iii) the hourly correlation between each time series. The operating scenarios are drawn from this larger set of representative days. The choice of the number of representative days per year is based on the result presented by [39]. Lara et al. conclude that for this data set and case-study, 4 representative days per year were sufficient to adequately represent the variability in load and capacity factor for the different regions.

In order to further simplify the transmission model, the ”truck-route” representation is adopted, which assumes that the flow in each line can be determined by an energy balance between nodes. This approximation ignores Kirchhoff’s voltage law, which dictates that the power will flow along the path of least impedance. We also assume that the transmission lines have a maximum capacity, and that transmission expansion is not considered. Additionally, the transmission losses are characterized by a fraction loss per mile, and are not endogeneously calculated.

2.2 MSIP model

Assuming the data for this process is uncertain and evolves according to a stochastic process, the MSIP model can be concisely formulated as in (1), following the similar notation as [89].

$$\min_{(x_1,y_1)\in\chi_1}\left\{f_1(x_1,y_1)+\mathbb{E}_{\bar{\xi}_{[2,\Gamma]}|\xi_{[1,1]}}\left[\min_{(x_2,y_2)\in\chi_2(x_1,\xi_2)}\left\{f_2(x_2,y_2,\xi_2)+\right.\right.\right.$$
$$\left.\left.\left.\ldots+\mathbb{E}_{\bar{\xi}_{[\Gamma,\Gamma]}|\xi_{[1,\Gamma-1]}}\left[\min_{(x_\Gamma,y_\Gamma)\in\chi_\Gamma(x_{\Gamma-1},\xi_\Gamma)}\left\{f_\Gamma(x_\Gamma,y_\Gamma,\xi_\Gamma)\right\}\right]\right\}\right]\right\} \tag{1}$$

where $\gamma\in\{1,\ldots,\Gamma\}$ is the set of stages and Γ is the last stage; x_γ is the set of state variables that link different stages; y_γ is the set of local variables that do not depend on the decision of previous stages and is only contained in the subproblem at stage γ. In the context of this GEP problem, the state variables are the number of generators in cluster i of region r that are operational in year t and number of storage units in cluster j of region r that are operational in year t. State variables are mixed-integer as the number of active generators is forced to be integer for thermal units but it is allowed to be fractional for renewable generators and storage units (for details see Appendix A). It is important to highlight that while decision stages are ordered in time, individual stages can have one or multiple time-periods t within it.

$\chi_\gamma(x_{\gamma-1},\xi_\gamma)$ is the feasible region of the stage γ, which depends on the decisions in stage $\gamma-1$ and the uncertainty realization ξ_γ in stage γ. $\bar{\xi}[\gamma,\gamma']$ denotes a sequence of random data vectors corresponding to stages γ through γ' and $\xi[1,\gamma-1]$ denotes a specific realization of this sequence of random vectors from stage 1 to stage $\gamma-1$. $\mathbb{E}_{\bar{\xi}[\gamma,\Gamma]|\xi[1,\gamma-1]}$ denotes the expectation operation in stage γ with respect to the conditional distribution of $\bar{\xi}[\gamma,\Gamma]$ given realization $\xi[1,\gamma-1]$ in stage $\gamma-1$.

This stochastic process has a finite number of realizations in the form of a scenario tree $\mathscr{T}$, with Γ stages and a set of nodes in each stage denoted by S_γ. Each node n in stage $\gamma>1$ has a unique parent node $P(n)$ in stage $\gamma-1$. The stage containing node n is denoted by $\gamma(n)$. The set of children nodes of a node n is denoted by $C(n)$, such that if $n\in S_\gamma$ and $m\in C(n)$, then $m\in S_{\gamma+1}$. The set of nodes on the unique path from origin node 1 to node n, including the latter, is denoted by $Path(n)$. A node $n\in S_\gamma$ represents a state of the system in stage γ and corresponds to the sequence of realizations $\{\xi_m\}_{m\in Path(n)}$. The probability of node n to happen, which is the probability of realization of the sequence $\{\xi_m\}_{m\in Path(n)}$, is denoted $prob_n$. For a node in the last stage of the tree, $n\in S_\Gamma$, the sequence of realizations $\{\xi_m\}_{m\in Path(n)}$ is called a scenario $sc\in SC$ and the set of nodes n that are part of this scenario sc are denoted by S_{sc}. Therefore, the *extensive form* (also known as deterministic equivalent) of (1) can be formulated as:

$$\mathscr{P}:\min_{(x_n,y_n)}\left\{\sum_{n\in\mathscr{T}} prob_n\cdot f_n(x_n,y_n)\mid(x_{P(n)},x_n.y_n)\in\chi_n\ \forall\ n\in\mathscr{T}\right\} \tag{2}$$

A summary of the notation of main parts of the scenario tree is shown in Figure 1. The detailed MSIP formulation is described in Appendix A.

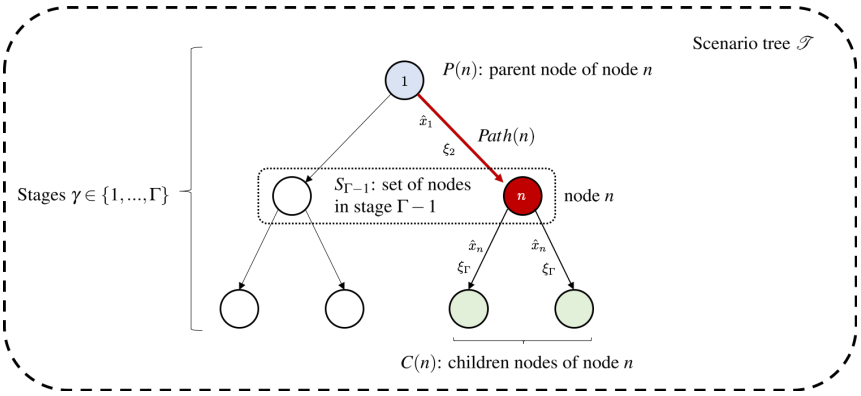

Fig. 1 Summary of the notation for scenario tree $\mathscr{T}$

2.3 Scenario tree generation

In our framework we capture both long-scale strategic and short-scale operating uncertainties. The strategic uncertainties occur in the same time scale as the investment decisions (i.e., yearly), while the operating uncertainties are captured through different representative days' profiles of hourly load demand and renewable availability between investment decisions.

Our GEP problem can be represented with the multi-horizon framework proposed by [34] and used by [44]. This methodology represents strategic and operating uncertainties separately based on the observation that strategic decisions typically do not depend directly on any particular operational scenario, implying that it is enough to branch only between strategic stages, and the operational decisions can be seen as embedded into (or attached to) their respective strategic nodes. However, multi-horizon representation is not needed as we assume stage-wise independence in the scenario tree (see Section 3.2.2 for more details).

For a problem with Ξ^{s} strategic realizations per stage and Ξ^{o} operational realizations per stage leads to a problem with a total number of scenarios of $\overline{SC} = (\Xi^{\mathrm{s}}\Xi^{\mathrm{o}})^{\Gamma-1}$. Figure 2 shows the standard and the recombining representations of the scenario tree $\mathscr{T}$. ~~They are~~ A recombining scenario tree is one in which for any two nodes n and n' in S_t, the set of children nodes $C(n)$ and $C(n')$ are defined by identical data and probabilities. Therefore, the standard tree and the recombining tree are equivalent in this context of stage-wise independence, when the uncertainty is memory-less.

This scenario tree has $\Xi^{\mathrm{s}} = 3$ strategic realizations per stage and $\Xi^{\mathrm{o}} = 2$ operational realizations per stage, hence, it has a total of $\overline{SC} = (6)^{\Gamma-1}$ scenarios.

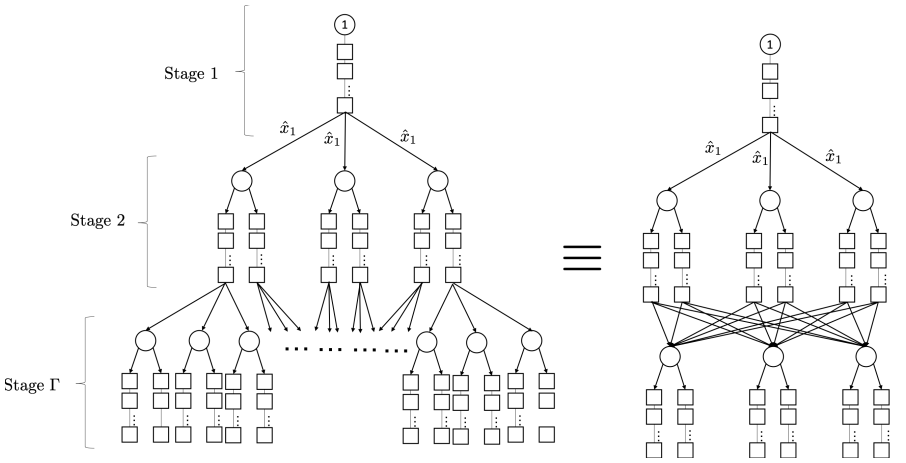

Fig. 2 Standard (left) and recombining (right) representations of the scenario tree $\mathscr{T}$ with both strategic and operating uncertainties. The circles represent the strategic decisions and while the squares represent the operating decisions.

3 SDDiP Decomposition

As mentioned in Section 1, multistage stochastic programming models grow exponentially with the number of stages, leading to a large multi-scale problem that quickly becomes intractable. Formulation 1 exploits the nested structure in this MSIP problem. Therefore, we use Stochastic Dual Dynamic Integer Programming (SDDiP) because it can take advantage this nested structure in the problem.

Birge was the first to apply Benders Decomposition on a nested fashion to solve Multistage Stochastic Linear Programming (MSLP) models, in 1985 [6]. A few years later, Pereira et al. [66] re-explained Nested Benders Decomposition using dynamic programming notation, incorporated scenario sampling into the algorithm, and showed that convergence can be significantly improved under the assumption of stage-wise independence, calling this new decomposition method Stochastic Dual Dynamic Programming (SDDP). Both Nested Benders Decomposition and SDDP convergence [6] and almost-sure convergence [67] proofs, respectively, rely on the fact that cost-to-go functions in MSLP models are piece-wise linear and convex. Therefore, both decomposition methods had their application limited to convex multistage stochastic programming models (and were mostly used for linear programming models).

Recently, there have been research efforts on extending Nested Benders/SDDP to integer and mixed-integer stochastic programming models. Cerisola et al. [10] propose a variant of Benders Decomposition for multistage stochastic integer programming and apply it to the stochastic unit commitment problem. Thome et al. [82] introduce an extension of the SDDP framework by using Lagrangean Relaxation to convexify the recourse function applied to nonconvex hydrothermal operation planning. Zou et al. [89] propose a valid Stochastic Dual Dynamic Integer Programming (SDDiP) algorithm for MSIP with binary state variables, and prove that for some of the cuts presented the algorithm converges in a finite number of steps. Lara et al. [39]

show that the cuts presented by Zou et al. [89] are still valid for problems with mixed-integer state variables. However, finite convergence is not guaranteed, i.e., there may be a duality gap.

The SDDiP algorithm consists of breaking down the scenario tree by nodes, and solving it iteratively in a forward and backward fashion until the optimality tolerance ε is satisfied, as shown in Figure 3. The Forward Pass yields a statistical upper bound, while the Backward Pass, which generates cuts from the relaxed subproblems to outer approximate the cost-to-go function, provides a lower bound. New cuts are added in the Backward Pass of each iteration k, and are kept in the following Forward Pass, until the difference between the upper and lower bounds is less then a pre-specified tolerance.

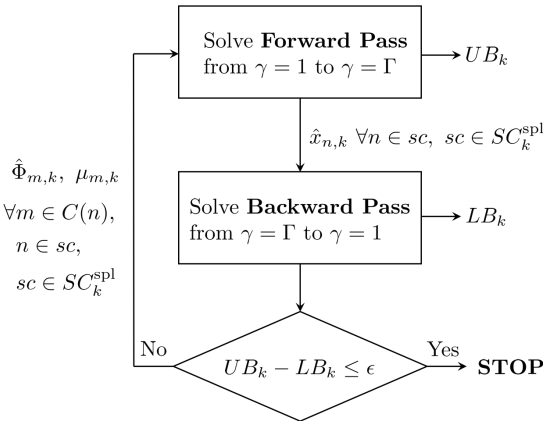

Fig. 3 Steps at iteration k of the SDDiP algorithm, where $\hat{\Phi}_{m,k}$ and $\mu_{m,k}$ are the coefficients of the cuts, sc is a scenario and SC^{spl_k} is the set of scenarios sampled in iteration k.

To be able to decompose the MSIP problem by node, we make copies of the state variables x_n. These new auxiliary variables, z_n, are used to equate to the parent node's state and make sure that when solving node n we are continuing from the state of the system at the end of its parent node. z_n is, however, relaxed to be a continuous variable within the same bounds as x_n.

Going back to our MSIP problem, the node subproblems of (2) can be formulated as follows $\forall\, n \in \mathscr{T}, k \in \mathscr{K}$:

$$
\begin{aligned}
\mathscr{P}_{n,k} : \Phi_{n,k}(\hat{x}_{P(n),k}, \phi_{n,k}) := \quad & \min_{(x_n, y_n)} \quad f_n(x_n, y_n) + \sum_{m \in C(n)} q_{nm} \phi_{m,k}(\hat{x}_{n,k}) \\
& s.t. \qquad (z_n, x_n . y_n) \in \chi_n \\
& \qquad\quad z_n = \hat{x}_{P(n),k} \quad \leftarrow \mu_{n,k} \in \mathbb{R}^{\ell} \\
& \qquad\quad x_n \in \mathbb{Z}_+^{\ell_1} \times \mathbb{R}_+^{\ell_2},\ \ y_n \in \mathbb{Z}_+^{o_1} \times \mathbb{R}_+^{o_2},\ \ z_n \in \mathbb{R}^{\ell}
\end{aligned}
\tag{3}
$$

where $\ell = \ell_1 + \ell_2$, $o = o_1 + o_2$; $\mathscr{K}$ is the set of iterations; and $q_{nm} := prob_m / prob_n$ is the conditional probability of transitioning from node n to node m for $m \in \mathscr{T} \setminus \{1\}$ and $n = P(m)$.

The approximate expected cost-to-go function, $\phi_{n,k}(\cdot)$, is defined as:

$$\phi_{n,k}(\hat{x}_{n,k}) := \min_{x_n, \alpha_n} \left\{ \alpha_n : \ \alpha_n \geq \sum_{m \in C(n)} q_{nm} \cdot \left(\hat{\Phi}_{m,k'} + \mu_{m,k'}^{\intercal} (\hat{x}_{n,k'} - x_n) \right) \ \forall \, k' \in \mathscr{K} | k' < k \right\} \tag{4}$$

3.1 Forward Step

The Forward Step for the SDDiP is very similar to the Forward step presented by [39], but now applied to the scenario tree with incorporated scenario sampling.

As mentioned before, the purpose of the Forward Pass in SDDiP is to generate a statistical upper bound to the solution of the MSIP for the entire scenario tree $\mathscr{T}$. It accomplishes this by randomly sampling a subset of the scenarios in the tree $\mathscr{T}$ - SC_k^{spl}, and for each stage $\gamma \in \{1, ..., \Gamma\}$ solving the nodes n in $S(\gamma)$ if they are also part of the sampled scenarios. By solving node n, the algorithm implements the optimal decisions for this node considering its uncertainty realization and previous $Path(n)$, and passes the current state of the system forward to its children node $m \in C(n)$ if m is also part of the sampled scenarios. This process is repeated up until all sampled scenarios are fully solved (until last stage Γ) and we have a total minimum cost for each of those scenarios.

The Forward Pass with scenario sampling is shown in Figure 4 for both the standard representation and the recombining scenario tree representation. For our case in which we assume stage-wise independence, both representations are equivalent.

The problem is assumed to have complete continuous recourse, which means that for any value of state variable (i.e., linking variable) and local integer variables, there are values for the continuous local variables such that the solution is feasible. This assumption is valid since feasibility can be achieved by adding nonnegative slack variables and penalizing them in the objective function.

The statistical upper bound, UB_k, is calculated in the Forward Pass as in (5).

$$UB_k = \bar{\mu}_k + z_{\alpha/2} \cdot \frac{\sigma_k}{\sqrt{N^{\text{spl}}}} \qquad \forall \, k \tag{5}$$

where $\bar{\mu}_k$ is the mean total cost over the sampled scenarios in iteration k, σ_k is its standard deviation, N^{spl} is the number of scenarios sampled in each iteration, and $z_{\alpha/2}$ is the z-score to assure a certain confidence interval (for example, for a 95% confidence interval, $z_{\alpha/2} = 1.96$).

The total cost of a scenario sc is as follows.

$$\Phi_{sc,k}^{\text{tot}} = \sum_{\gamma \in \{1,...,\Gamma\}} \sum_{n \in S_\gamma \cap n \in S_{sc}} \left(\hat{\Phi}_{n,k} - \hat{\alpha}_n \right) \qquad \forall sc \in SC_k^{\text{spl}} \tag{6}$$

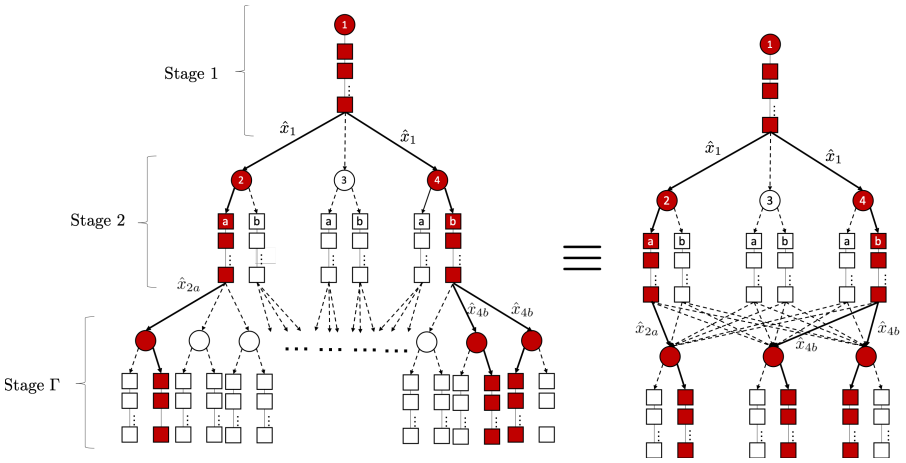

Fig. 4 Forward Pass with scenario sampling for both the standard and the recombining scenario tree representations. The nodes highlighted are the ones that are part of the sampled scenarios, and the continuous arrows show the paths of the sampled scenarios.

and $\bar{\mu}_k$ and σ_k are defined in (7) and (8), respectively.

$$\bar{\mu}_k = \frac{1}{N^{\text{spl}}} \sum_{sc \in SC_k^{\text{spl}}} \Phi_{sc,k} \qquad \forall k \in \mathscr{K} \tag{7}$$

$$(\sigma_k)^2 = \frac{1}{N^{\text{spl}} - 1} \sum_{sc \in SC_k^{\text{spl}}} \left(\Phi_{sc,k} - \bar{\mu}_k\right)^2 \qquad \forall k \in \mathscr{K} \tag{8}$$

It is important to note that the upper bound, UB_k, obtained by the SDDiP is only a statistical upper bound. Its validity is guaranteed with certain probability provided that N^{spl} is not too small. However, regardless of the size of N^{spl}, it is possible that the upper bound is smaller than the valid lower bound evaluated in the backward step. To avoid this issue, alternative stopping criteria are reported in the literature [55,79,8]. We, however, ~~are using~~ use the standard stopping criteria of $UB_k - LB_k \leq \varepsilon$, where ε is the allowed optimality tolerance. This is the same stopping criteria as the one used by [89], but we acknowledge it may lead to an early termination of the algorithm.

3.2 Backward Step

After solving the Forward Pass, the next step is the Backward Pass, and its purpose is to generate cuts that outer approximate the cost-to-go function. The Backward Pass consists of solving the subproblems from the last to the first stage, so the solutions of future stages can be used to generate cuts and provide approximations to the cost-to-go functions within the planning horizon.

Since our MSIP has mixed-integer recourse (i.e., the state variables x_n are mixed-integer), instead of solving the original subproblems we have to solve a relaxation that is convex in the subspace of the state variables in order to generate a valid cut. This

relaxation can be the linear programming relaxation or the Lagrangean relaxation of the subproblem $\mathscr{P}_{n,k}$ given by (3). Depending on the type of relaxation solved, a different type of cut is added in the Backward Pass to approximate the cost-to-go function.

In this step, instead of only having to solve the nodes that are part of the scenarios sampled in iteration k, SC_k^{spl}, we have to solve the subproblem of all the children nodes $C(n)$ of the nodes n that are part of the sampled scenarios $n \in S_{sc}, sc \in SC_k^{\text{spl}}$. Consequently, we solve a total of $(N^{\text{spl}} \cdot \Xi^{\text{s}} \cdot \Xi^{\text{o}})$ subproblems, as can be seen in Figure 5. The solution of this extra set of nodes is necessary to be able to generate the cuts that approximate the cost-to-go function, which takes an weighted average of the coefficients coming from the solution of the subproblem of the children nodes based on the probabilities of the uncertainty realizations.

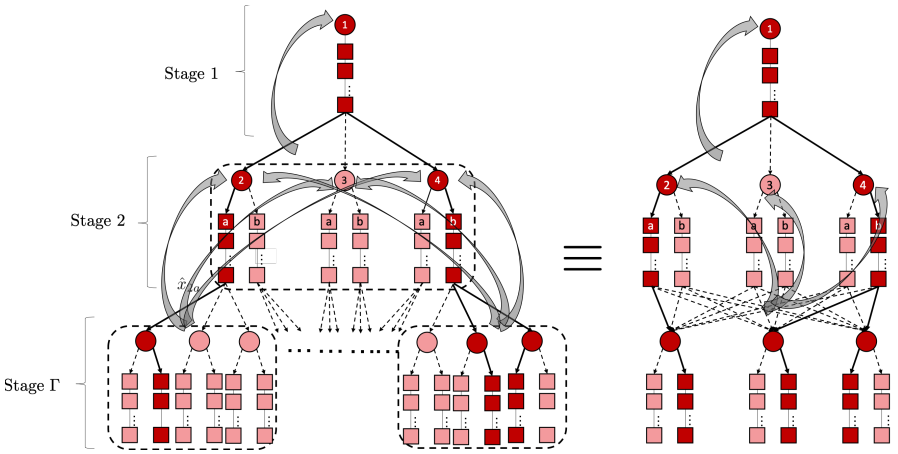

Fig. 5 Backward Pass with scenario sampling and stage-wise independence for the standard (left) and recombining (right) scenario tree representations. The nodes in dark color are the ones that are part of the sampled scenarios, and the nodes in lighter color are the nodes that are not part of the sampled scenarios but were solved because they are children nodes of the sampled nodes.

The lower bound, LB_k, is calculated in the Backward Pass as in (9). It is easy to see that the relaxed solution of the root node $n = 1$ is a lower bound to the total cost since it only has a subset of the original constraints of the original problem.

$$LB_k = \hat{\Phi}_{1,k} \qquad \forall\, k \in \mathscr{K} \tag{9}$$

3.2.1 Possible cuts to approximate the cost-to-go function

The choice of cuts directly impacts the performance of the algorithm as some cuts are tighter and more/less computationally expensive to generate than the others. The Benders cut, Strengthened Benders cut and Lagrangean cut were first proposed by Zou et al. [89] for MSIP with binary recourse, and Lara et al. [39] proved their validity for models with mixed-integer recourse. However, it is important to highlight that the

SDDiP algorithm does not have guaranteed finite convergence if the formulation has mixed-integer recourse and the Backward Pass uses any of the following cuts, which means that there can be a duality gap.

Benders cut The first option trivially comes from SDDP Benders cut. The Benders cut's coefficients are obtained from the solution of the linear relaxation (LP) of $\mathscr{P}_{n,k}$ in (3), and is formulated as follows.

$$\alpha_n \geq \sum_{m \in C(n)} q_{nm} \cdot \left(\hat{\Phi}^{\mathrm{LP}}_{m,k'} + \mu^{\mathrm{LP}\ \intercal}_{m,k'} (\hat{x}_{n,k'} - x_n) \right) \qquad \forall\, k' \in \mathscr{K} | k' < k \tag{10}$$

This is the weakest of the possible cuts, but it has the advantage of being easily and quickly computed. As shown by [39], the Benders cut performs very well for this GEP problem, since it has a tight linear relaxation. For certain multistage capacity planning problems with integer recourse, there is evidence that Benders cuts alone are sufficient for reducing the optimality gap to zero as the number of stages increases (see, e.g., [29], Corollaries 1 and 2).

Lagrangean cut The subproblem $\mathscr{P}_{n,k}$ can also be convexified by considering its Lagrangean relaxation, which yields the convex hull of the nonlinking constraints [19]. In our case this is done by dualizing the linking equalities in (3) and penalizing their violation in the objective function by the vector of Lagrange multipliers, $\mu_{n,k}$. The closer the Lagrange multipliers are to their optimal value, the tighter the approximation is, and the stronger the cuts generated by theses multipliers are. Therefore, the Lagrangean cut uses the coefficients obtained by the Lagrangean Dual (LD) problem and is formulated as follows.

$$\alpha_n \geq \sum_{m \in C(n)} q_{nm} \cdot \left(\hat{\Phi}^{\mathrm{LD}}_{m,k'} + \mu^{\mathrm{LD}\ \intercal}_{m,k'} (\hat{x}_{n,k'} - x_n) \right) \qquad \forall\, k' \in \mathscr{K} | k' < k \tag{11}$$

For more details on the Lagrangean cut, see [89, 39].

Strengthened Benders cut In order to mitigate potential performance issues, Zou et al. [89] proposed the Strengthened Benders cut, which is a compromise between Benders and Lagrangean cuts. Its generation is similar to the Lagrangean cut, but it does not use the subgradient method to improve the multipliers. Instead, it uses the coefficients from the first Lagrangean relaxation (LR) solved after the initialization of the multipliers using LP relaxation and is formulated as follows.

$$\alpha_n \geq \sum_{m \in C(n)} q_{nm} \cdot \left(\hat{\Phi}^{\mathrm{LR}}_{m,k'} + \mu^{\mathrm{LP}\ \intercal}_{m,k'} (\hat{x}_{n,k'} - x_n) \right) \qquad \forall\, k' \in \mathscr{K} | k' < k \tag{12}$$

For more details on the Strengthened Benders cut, see [89, 39].

Due to the computational expense of computing the Lagrangean and the Strengthened Benders cuts, and the computational evidence in [39] that the Benders cuts are likely sufficinet for this GEP problem, we select the Benders cuts to be used in our computational experiments shown in Section 4.

3.2.2 Stage-wise independence and Cut Sharing

In multistage problems, if the stochastic process and the constructed scenario tree is stage-wise independent, i.e., for any two nodes n and n' in S_t the set of children nodes $C(n)$ and $C(n')$ are defined by identical data and conditional probabilities, then the cost-to-go functions do not depend on the current scenario. This means that the value functions and expected cost-to-go functions depend only on the stage rather than the nodes, $\Phi_n(\cdot) \equiv \Phi_\gamma(\cdot)\ \forall\ n \in S_\gamma$, and the cuts generated for a particular scenario are also valid for any other scenario at the same stage [31].

The SDDiP relies on the stage-wise independence assumption and the ability to share cuts among different nodes in the same stage to avoid the combinatorial explosion and the "curse of dimensionality" [66]. Therefore, it provides a practical solution for solving real-world applications of MSIP on very large scenario trees without the need for scenario reduction methods.

The Backward Pass in the SDDiP algorithm works similarly as the one in a Stochastic Nested Benders decomposition with sampling. The only difference is that because we have stage-wise independence, the cuts generated are added to all the nodes in the previous stage instead of only to the parent node. The Backward Pass with stage-wise independence assumption is shown is Figure 5.

The stagewise independence assumption is reasonable for the operating uncertainties considered in our GEP problem (i.e., different profiles for representative days) since the realization of the solar incidence, wind speed and load profile in ~~one day~~ stage t (year t) has little influence in the ~~next day~~realizations of the next stage (year $t+1$), especially if we are only including a few representative days a year. These representative days are not sequential, so there are no operational decisions that connect the days, which is a reasonable assumption considering that midnight is far away from the extreme ramps of the duck curve [13].

Regarding the strategic uncertainties, the stage-wise independence assumption is adequate for natural gas price uncertainty as the prices can go up and down without a clear influence of the realization in the year before. In the case of carbon tax uncertainty, this assumption may be a stretch as it is unlikely that the carbon tax would wildly vary between two consecutive years. However, even though carbon tax stage-wise independence generates an exceptionally uncertain mode, we still believe that in face of the the political uncertainty and changes in administration over the planning horizon this is a case worth examining.

Additionally, there has been some work on how to extend the use of SDDP/SDDiP for certain types of interstage dependency [31,79,72,45], which would be useful to capture GEP uncertainties that are not well-captured by the stage-wise independence assumptions (e.g., learning rate of new generation and storage technologies and peak load). For the case of strategic uncertainty with stage-wise dependence and operating uncertainty with stage-wise independence, one can use the framework proposed by Rebennack [75], which combines SDDP with the sampling-based stochastic nested Benders decomposition approach. These extensions are left for future investigation.

3.3 Parallelization Scheme

In the SDDiP framework the subproblems of the nodes n within the stage γ, $n \in S_\gamma$, are independent from each other. Hence, SDDiP is well suited for parallel processing. The algorithm is not, however, trivially parallel since synchronization is required to share the Benders cuts generated with all the nodes in the previous stage $\gamma - 1$. Therefore, a well-thought parallelization strategy is required in order to obtain an efficient parallel solution.

There has been some effort in the literature to propose the optimal parallelization scheme for SDDP in order to avoid synchronization steps as much as possible [69, 27]. We do not claim that our parallelization scheme is optimal, but based on the results that we obtained it seems to be adequate for our problem.

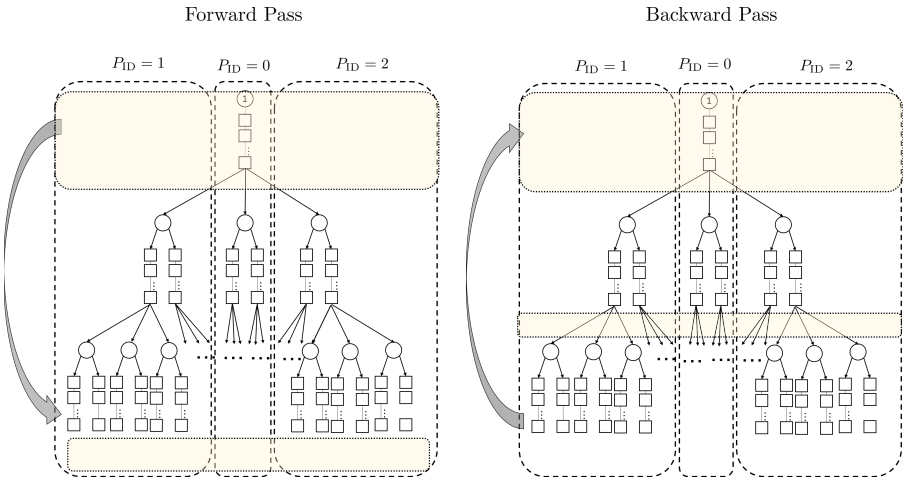

Fig. 6 Parallelization scheme used for the SDDiP algorithm. The dashed lines show which process is in charge of which node, and the highlighted areas show the synchronization points both in the Forward and Backward Passes.

We use PyMP [41] for the parallelization, which is a Python package based on OpenMP [11]. We first equally divide the nodes in the tree among the processes, and then enter the parallel context and have each process generate the subproblems for those nodes assigned to it. After that, we start the Forward Pass and randomly select the sampled scenarios to be solved for each of those processes, $sc \in SC_k^{\mathrm{spl,pid}}$, such that the number of processes N^{pids} times the number of sampled scenarios by processes $N^{\mathrm{spl.pid}}$ equals the total number of sampled scenarios per iteration k: $N^{\mathrm{pids}} \cdot N^{\mathrm{spl,pid}} = N^{\mathrm{spl}}$. We then solve the subproblems of all the nodes that are part of the sampled scenarios, storing the results of the state variables as shared dictionaries among the processes. Note that as the nodes are statistically assigned to processes there is the potential for load imbalance due to both the random sample of scenarios and because of variance in the time to solve each MILP subproblem. Addressing this potential scaling issue is left for future work.

When all processes reach the end of the Forward Step, there is a synchronization step to gather all the optimal values before the the first process calculates the upper bound UB_k and distributes it to the other processes. In the Backward Pass there are synchronization steps after every stage to distribute cuts generated at that stage to all nodes in the previous stage. In the end only the first process calculates the lower bound LB_k and check if the optimality tolerance ε is satisfied. The parallelization scheme and the synchronization points for both the Forward and Backward Passes are shown in Figure 6.

Our implementation of the parallel SDDiP algorithm for this GEP problem can be found in [38].

4 Case study: ERCOT region

We test the proposed MSIP formulation and SDDiP algorithm for a case study approximating the Texas Interconnection, a power grid that covers most of the state of Texas and is managed by the Electric Reliability Council of Texas (ERCOT). This case study is based on the deterministic case study presented by [39], with the addition of operational and strategic uncertainties, and the option of adding storage units.

Within the ERCOT covered area, we consider four geographical regions: Northeast, West, Coastal and South. We also include a fifth region, Panhandle, which is technically outside the ERCOT limits but due to its renewable generation potential, it supplies electricity to the ERCOT regions. Thus, Panhandle is considered a zone with zero load demand, i.e., it is only a supplier, not a consumer. The regions are shown in Figure 7. For more details on the sources of the data used, see [39].

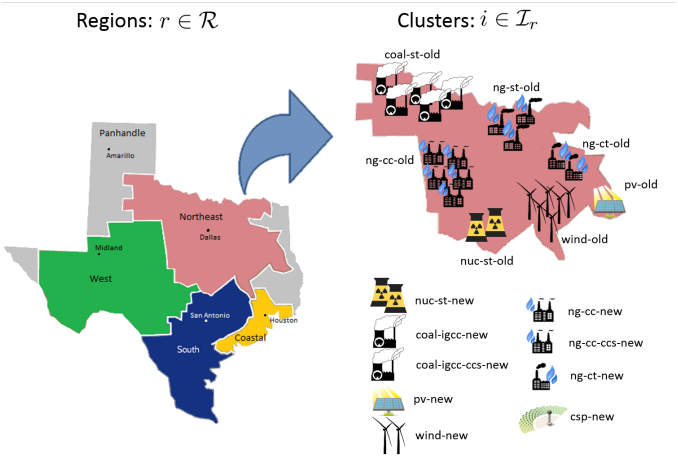

Fig. 7 Model representation of regions and clusters [39]

We consider 3 types of utility batteries: lithium-ion, lead-acid and flow batteries, for which we use the capital cost forecast provided by Schmidt et al. [78] and the technical information provided by Luo et al. [50], the same sources used by [40].

4.1 Reference case: all energy sources included

For each of the regions, we use load and capacity factor profiles with an hourly resolution. Representative days are constructed using a k-means clustering algorithm and 2004-2010 zonal load and renewables profiles, as explained in [39]. 8 clusters are constructed to find the 8 most representative days, and these days are split in 2 scenarios which correspond to 2 realizations of the operational uncertainty per stage, with 4 representative days each. We can assume that the operational uncertainty satisfies the stage-wise independence assumption discussed in Section 3.2.2 and follows a uniform distribution.

Additionally, we consider that the natural gas fuel price is uncertain and has 3 realizations per stage. We assume that this fuel price is stage-wise independent and follows a uniform distribution. The realizations were built using the minimum, median and maximum value corresponding to the scenarios presented in the EIA Annual Energy Outlook 2019 [17]. We assume that the coal and uranium prices are deterministic since they exhibit considerably less variation compared to the natural gas price.

Our computational tests are performed on a MacBook Pro with 2.3GHz quad-core 8th-generation Intel Core i5 processor, with 8GB of RAM, running on macOS 10.14 Mojave. We implement the SDDiP algorithm in Python 3.6.6 and Pyomo 5.6.1, and solve the LPs and MILPs of each node of the the scenario tree using Gurobi version 8.0.1 [26]. We allow a total number of 3 parallel processes, sample 15 scenarios per iteration, impose a 95% confidence interval in the statistical upper bound ($z_{\alpha/2} = 1.96$), and consider that the algorithm converges if it reaches an optimality gap of less than or equal to 1%.

We first test the parallel efficiency of our algorithm by solving the same 5-stage problem both sequentially and in parallel (we fix the random seed to avoid the stochasticity in the solution by the random sampling of scenarios per iteration). The solution time for the serial implementation is $t_s = 13,950$ seconds and for the parallel implementation with 3 processes is $t_p = 6,131$. Therefore, our parallel SDDiP algorithm has 76% efficiency.

To test the capabilities of our algorithm, we solve the problem for: (i) 5 stages (5 years, 1 year per stage); (ii) 10 stages (10 years, 1 year per stage); and (iii) 15 stages (15 years, 1 year per stage). In all of them, the subproblem (node) size before cuts is 50,042 constraints, 13,746 integer variables, and 22,755 continuous variables. The size of the extensive form (deterministic equivalent) and performance of the SDDiP algorithm for all cases are reported in Table 1.

The extensive forms of all three cases are massive, with up to quadrillions of variables and constraints in the 15-stage case. Considering the size of the models, all cases were solved in a reasonable amounts of time: 1.7 hours, 2.3 hours and 23.9 hours, respectively. Due to memory limitations, it is fair to say that at least the 10-stage and 15-stage cases would not be solvable in a personal laptop or desktop without decomposing the model, or if the model is decomposed by a scenario-based approach without scenario reduction techniques. Additionally, the convergence time is considerably less than similar sized GEP problems with multistage stochastic programming formulations reported in the literature that use scenario-based decomposition (see

Table 1 Size of the problem and SDDiP performance for scenario trees with different numbers of stages

	5 stages	10 stages	15 stages
Number of scenarios	$1,296$	1.01×10^7	7.84×10^{10}
Number of nodes in the scenario tree	$1,555$	1.21×10^7	9.40×10^{10}
Number of constraints (extensive form)	7.78×10^7	6.05×10^{11}	4.71×10^{15}
Number of integer variables (extensive form)	2.14×10^7	1.66×10^{11}	1.29×10^{15}
Number of continuous variables (extensive form)	3.54×10^7	2.75×10^{11}	2.14×10^{15}
Wall-clock time [s]	$6,131$	$8,146$	$86,049$
Upper bound [$ billion]	51.69	91.87	122.40
Lower bound [$ billion]	51.69	91.27	121.72
Optimality gap [%]	8.57×10^{-4}	0.66	0.56

[44]). These results show how powerful and useful SDDiP can be for practical large-scale MSIP models.

An important question that can now be explored is how impactful the length of the planning horizon and the number of stages are in the optimal "here-and-now" first stage investment decisions. For this reference case, even though there are marginal differences between the optimal first-stage decisions depending on the number of stages (e.g. in the 15-stage solution the optimization adds 2 PV solar units in the first year while the 5-stage and 10-stage solutions do not), the results are extremely similar (2.27 GW, 2.21 GW, 2.40 GW of natural gas generation capacity are added in year one, respectively), indicating that solving a 5-stage problem would be sufficient.

This result is not surprising considering that the ERCOT system is mature and stable, and therefore we would not expect to see drastic changes without a corresponding drastic impulse on the system. Accordingly, the optimization chooses to install as few generators as possible in this first year, and wait until some of the uncertainty is realized to make future decisions.

4.2 *No nuclear* case

The value of stochastic programming with multiple decision stages becomes more accentuated in power systems where there is a need for quick and significant expansion, which is the case in some developing countries (e.g., India's electricity demand is expected to double over the coming decade [30]).

Therefore, we solve a hypothetical case in which ERCOT decides to immediately retire nuclear power. This is an interesting analysis considering the declining profits and scheduled retirements of nuclear plants in the United States [83]. The nuclear reactors represent 5% of the initial ERCOT generation capacity in the original data set. Thus, by imposing their immediate retirement, the optimization is forced to make significant expansion decisions in the first stage (first year).

We first solve the *no nuclear* case study with the same scenario tree as before, i.e. 2 realizations of operational uncertainty per stage, 3 realizations of natural gas price per stage, and 15 stages. The optimal ERCOT generation capacity by source in the

first year (first stage) for the reference case (all sources included and natural gas price uncertainty) and the *no nuclear* case with natural gas price uncertainty are shown in Figure 8. The results show that the nuclear plants were fully replaced by natural gas combined-cycle plants.

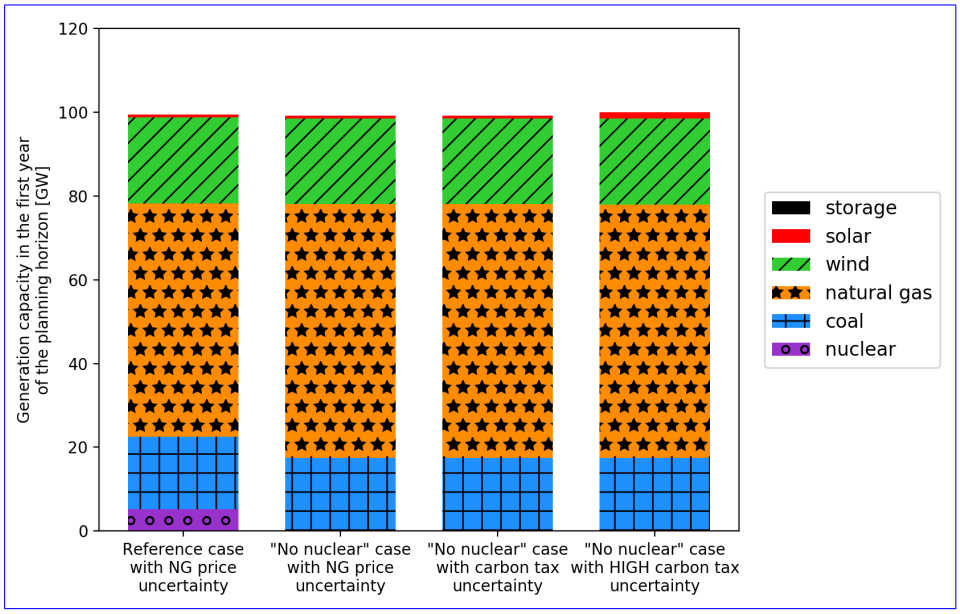

Fig. 8 ERCOT generation capacity by source in the first year (origin node) for the reference case (all sources included and natural gas price uncertainty) and the *no nuclear* case with natural gas price uncertainty, carbon tax uncertainty, and high carbon tax uncertainty.

A potential issue that would arise if all nuclear reactors are replaced by natural gas turbines is the increase of CO_2 emissions [83]. Therefore, we also solve the *no nuclear* case study with carbon tax uncertainty. We consider 2 realizations of operational uncertainty per stage (same as before), and 3 realizations of carbon tax price per stage such that the *no* realization of carbon tax equals \$0.0/tonne CO_2 for all stages, the *high* realization starts at \$10/tonne CO_2 at year 2 (stage 2) and increases linearly to \$150/tonne CO_2 at year 15 (stage 15), and the *medium* realization is the average between *low* and *high* realizations. We assume that the carbon tax prices are stage-wise independent and follow a uniform distribution. As mentioned in Section 3.2.2, the carbon tax stage-wise independence assumption generates an "exceptionally uncertain" model. The optimal ERCOT generation capacity by source in the first year (first stage) for the *no nuclear* case with carbon tax uncertainty is also shown on the right of Figure 8. The results show that even with the risk of having carbon tax fees in the future, the optimal "here-and-now" decision is to invest on new natural gas combined-cycle plants.

Additionally, we solve the *no nuclear* case study with a scenario tree that considers more extreme realizations of carbon tax: the *low* realization of carbon tax

equals \$0.0/tonne CO_2 for all stages (same as before), the *high* realization starts at \$100/tonne CO_2 at year 2 (stage 2) and increases linearly to \$500/tonne CO_2 at year 15 (stage 15), and the *medium* realization is the average between *low* and *high* realizations. We assume again that the carbon tax prices are stage-wise independent and follow a uniform distribution.

The reasoning behind solving the *no nuclear* case study for this more extreme scenario tree is to find out if natural gas will stop being the most attractive source if the variability between the realizations of carbon tax is higher. The optimal ERCOT generation capacity by source in the first year (first stage) for the *no nuclear* case with high carbon tax uncertainty is also shown on the right of Figure 8. The results show that the risk of having steep carbon tax fees in the future makes the optimization invest less in natural gas technologies in the first year (reduction of 0.58 GW in natural gas generation capacity), and more in renewable sources (increase of 0.74 GW and 0.04 GW in solar and wind generation capacity, respectively).

4.3 The value of stochastic solution

In order to evaluate the potential gain of solving this GEP as a MSIP model instead of a deterministic model, we solve a deterministic version of the *no nuclear* case study with high carbon tax uncertainty using the average of the carbon tax realizations. The comparison between the generation capacity by generation technology in the first year (first stage) for the MSIP formulation and the deterministic formulation is shown in Figure 9.

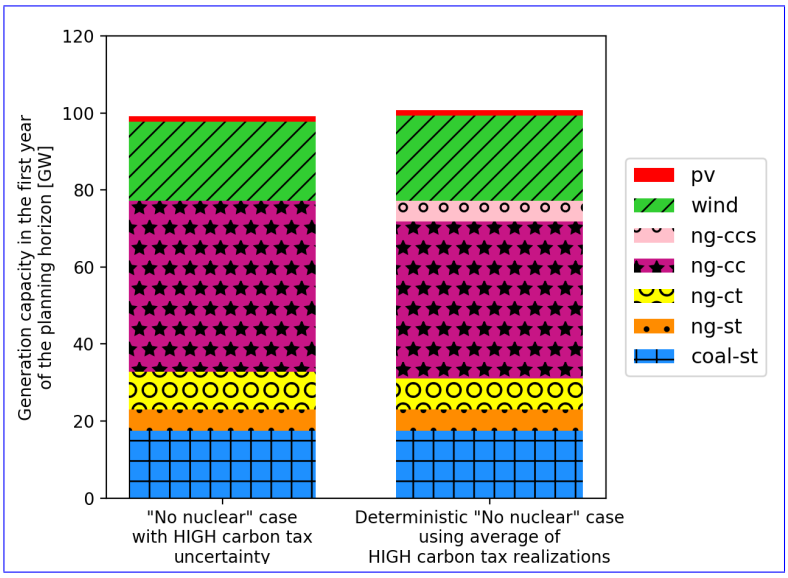

Fig. 9 ERCOT generation capacity by generation technology in the first year (origin node) for the *no nuclear* case with high carbon tax uncertainty and the deterministic solution using the carbon tax averages.

The results show that by assuming that carbon tax is a deterministic parameter, the optimization makes more conservative decisions and replaces some of the natural gas combined-cycle (ng-cc) and gas-fired combustion turbine (ng-ct) by natural gas combined-cycle with carbon capture (ng-cc-ccs) to avoid having to pay carbon tax for their emissions later. It also installs more wind turbines in the deterministic case than in the stochastic case. These results make sense because in the deterministic case the optimization is sure that there will be carbon tax in the future, while in the stochastic case there may be no carbon tax, a high carbon tax or a medium carbon tax, therefore it is better to wait and get more information about the carbon tax realization before investing in more expensive low-emission options.

To evaluate how well the deterministic first-stage solution would perform in our high carbon tax uncertainty scenario tree, we re-solve the MSIP formulation for the *no nuclear* case study with high carbon tax uncertainty fixing the investment decisions in the first stage to be the ones given by the deterministic solution. While the original stochastic solution gives an optimal expected value of $ 232.47 billion (0.28% optimality gap) the stochastic solution using the first-stage solution of the deterministic model gives an expected value of $ 234.65 billion (0.20% optimality gap), showing that by considering carbon tax an uncertain parameter the value of the stochastic solution is $2.18 billions, which is the savings one can achieve in the long term.

5 Conclusion

In this paper, we have proposed a multistage stochastic mixed-integer programming formulation to address the long-term generation expansion planning (GEP) under uncertainty including operating details ~~in~~ at the hourly level, storage options, and multiscale representation of uncertainty (strategic and operational). We ~~decompose the model~~ solve the problem using a parallelized SDDiP algorithm~~and show that this is a powerful framework for solving practical large-scale MSIP formulations,~~ , which is based on Nested Benders decomposition. By utilizing stage-based instead of scenario-based decomposition we were able to break down the proposed GEP model into smaller subproblems (one per node of the scenario tree), which had a great impact on its solution efficiency as the single-scenario formulation for this problem is already computationally demanding. Additionally, the assumption of stage-wise independence and the parallel processing were key for speeding up convergence and allowing the solution of ~~models with~~ instances of which the extensive form has up to quadrillions of variables and constraints in a personal computer. This shows how powerful the parallelized SDDiP framework can be for solving practical large-scale MSIP models. The stage-wise independence uncertainty, however, limits its usability and, depending on the case study, can generate a overly uncertain and potentially unrealistic scenario tree. Therefore, an important next step for this research would be to incorporate some of the work done in SDDP with stage dependency for the SDDiP framework.

We solved a hypothetical case study for the ERCOT region and show that for most of the scenario trees tested (with natural gas price and carbon tax uncertainty)

the first-stage decisions consist of investing in new natural gas plants, indicating the competitiveness of this source for this case study. We also ran a case study in which all nuclear reactors are immediately retired, and unless we consider steep values for the high realization of carbon tax the optimization still decides to invest in natural gas turbines in the first stage. Finally, we show the value of stochastic solution for the scenario tree with high carbon tax uncertainty, with a potential $ 2 billion reduction in cost in the long run.

As future work ~~it would~~ we would like to focus on the data and scenario generation, and perform a detailed out-of-sample simulation as a way to validate the importance of representing this model as multistage stochastic programming. We would also like to compare the result with those obtained from simplifications commonly used in the literature (e.g. representing this as a two-stage problem, or as multiple two-stage problems with rolling horizon) to evaluate the potential improvements of our proposed framework.

Additionally, it would also be interesting to consider the lead time of construction of the power plants as both a deterministic and an uncertain parameter to evaluate how this would impact in the planning strategy. Another addition to this work would be to evaluate and improve the parallel scalability of the proposed algorithm.

Acknowledgements We gratefully acknowledge funding from the U.S. Department of Energy, Office of Fossil Energys Crosscutting Research Program through the Institute for the Design of Advanced Energy Systems (IDAES). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

1. Albornoz, V.M., Benario, P., Rojas, M.E.: A two-stage stochastic integer programming model for a thermal power system expansion. International Transactions in Operational Research **11**(3), 243–257 (2004). DOI 10.1111/j.1475-3995.2004.00456.x
2. Babatunde, O.M., Munda, J.L., Hamam, Y.: Generation expansion planning: A survey. 2018 IEEE PES/IAS PowerAfrica pp. 307–312 (2018)
3. Baringo, L., Baringo, A.: A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. IEEE Transactions on Power Systems **33**(1), 792–802 (2018). DOI 10.1109/TPWRS.2017.2713486
4. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Adjustable robust solutions of uncertain linear programs. Mathematical Programming **99**(2), 351–376 (2004). DOI 10.1007/s10107-003-0454-y
5. Bertsimas, D., Sim, M.: The price of robustness. Operations Research **52**(1), 35–53 (2004). DOI 10.1287/opre.1030.0065
6. Birge, J.R.: Decomposition and partitioning methods for multistage stochastic linear programs. Operations Research **33**(5), 989–1007 (1985). DOI 10.1287/opre.33.5.989
7. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming, 2nd edn. Springer Publishing Company, Incorporated (2011)
8. Bruno, S., Ahmed, S., Shapiro, A., Street, A.: Risk neutral and risk averse approaches to multistage renewable investment planning under uncertainty. European Journal of Operational Research **250**(3), 979 – 989 (2016). DOI https://doi.org/10.1016/j.ejor.2015.10.013

9. Care, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. Operations Research Letters **24**(1), 37 – 45 (1999). DOI https://doi.org/10.1016/S0167-6377(98)00050-9
10. Cerisola, S., Baillo, A., Fernandez-Lopez, J.M., Ramos, A., Gollmer, R.: Stochastic Power Generation Unit Commitment in Electricity Markets: A Novel Formulation and a Comparison of Solution Methods. Operations Research **57**(1), 32–46 (2009). DOI 10.1287/opre.1080.0593
11. Chandra, R., Dagum, L., Kohr, D., Maydan, D., McDonald, J., Menon, R.: Parallel Programming in OpenMP. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2001)
12. Chen, C., Li, Y., Huang, G., Li, Y.: A robust optimization method for planning regional-scale electric power systems and managing carbon dioxide. International Journal of Electrical Power & Energy Systems **40**(1), 70 – 84 (2012). DOI https://doi.org/10.1016/j.ijepes.2012.02.007
13. Denholm, P., O'Connell, M., Brinkman, G., Jorgenson, J.: Overgeneration from solar energy in california. a field guide to the duck chart DOI 10.2172/1226167
14. Dentcheva, D., Römisch, W.: Optimal power generation under uncertainty via stochastic programming. In: K. Marti, P. Kall (eds.) Stochastic Programming Methods and Technical Applications, pp. 22–56. Springer Berlin Heidelberg, Berlin, Heidelberg (1998)
15. Diamant, A.: The Electric Generation Expansion Analysis System (EGEAS) Software. Tech. rep., Electric Power Research Institute (2017)
16. Ding, J., Somani, A.: A long-term investment planning model for mixed energy infrastructure integrated with renewable energy. In: 2010 IEEE Green Technologies Conference, pp. 1–10 (2010). DOI 10.1109/GREEN.2010.5453785
17. EIA: Annual Energy Outlook 2019. Tech. rep., U.S. Energy Information Administration (2019). URL https://www.eia.gov/outlooks/aeo/pdf/aeo2019.pdf
18. Flores-Quiroz, A., Palma-Behnke, R., Zakeri, G., Moreno, R.: A column generation approach for solving generation expansion planning problems with high renewable energy penetration. Electric Power Systems Research **136**, 232 – 241 (2016). DOI http://dx.doi.org/10.1016/j.epsr.2016.02.011
19. Frangioni, A.: "about lagrangian methods in integer optimization", journal="annals of operations research" **139**(1), 163–193 (2005). DOI 10.1007/s10479-005-3447-9
20. Gacitua, L., Gallegos, P., Henriquez-Auba, R., Lorca, A., Negrete-Pincetic, M., Olivares, D., Valenzuela, A., Wenzel, G.: "a comprehensive review on expansion planning: Models and tools for energy policy analysis". Renewable and Sustainable Energy Reviews **98**, 346–360 (2018). DOI https://doi.org/10.1016/j.rser.2018.08.043. URL http://www.sciencedirect.com/science/article/pii/S1364032118306269
21. Gandulfo, W., Gil, E., Aravena, I.: Generation capacity expansion planning under demand uncertainty using stochastic mixed-integer programming. In: 2014 IEEE PES General Meeting — Conference Exposition, pp. 1–5 (2014). DOI 10.1109/PESGM.2014.6939368
22. Gil, E., Aravena, I., Crdenas, R.: Generation capacity expansion planning under hydro uncertainty using stochastic mixed integer programming and scenario reduction. IEEE Transactions on Power Systems **30**(4), 1838–1847 (2015). DOI 10.1109/TPWRS.2014.2351374
23. Gorenstin, B.G., Campodonico, N.M., Costa, J.P., Pereira, M.V.F.: Power system expansion planning under uncertainty. IEEE Transactions on Power Systems **8**(1), 129–136 (1993). DOI 10.1109/59.221258
24. Grossmann, I.E., Apap, R.M., Calfa, B.A., Garca-Herreros, P., Zhang, Q.: Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. Computers & Chemical Engineering **91**, 3 – 14 (2016). DOI https://doi.org/10.1016/j.compchemeng.2016.03.002. 12th International Symposium on Process Systems Engineering & 25th European Symposium of Computer Aided Process Engineering (PSE-2015/ESCAPE-25), 31 May - 4 June 2015, Copenhagen, Denmark
25. Gupta, V., Grossmann, I.E.: A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties. Computers & Chemical Engineering **62**, 62 – 79 (2014). DOI https://doi.org/10.1016/j.compchemeng.2013.11.011
26. Gurobi Optimization, L.: Gurobi optimizer reference manual (2018). URL http://www.gurobi.com
27. Helseth, A., Braaten, H.: Efficient parallelization of the stochastic dual dynamic programming algorithm applied to hydropower scheduling. Energies **8**(12), 14287–14297 (2015). DOI 10.3390/en81212431
28. Heuberger, C.F., Rubin, E.S., Staffell, I., Shah, N., Dowell, N.M.: Power capacity expansion planning considering endogenous technology cost learning. Applied Energy **204**, 831 – 845 (2017). DOI https://doi.org/10.1016/j.apenergy.2017.07.075
29. Huang, K., Ahmed, S.: The value of multistage stochastic programming in capacity planning under uncertainty. Operations Research **57**(4), 893–904 (2009)

30. IEEFA: Indias Electricity Sector Transformation. Tech. rep., Institute for Energy Economics and Financial Analysis (2017). URL http://ieefa.org/wp-content/uploads/2017/11/India-Electricity-Sector-Transformation_Nov-2017-3.pdf
31. Infanger, G., Morton, D.P.: Cut sharing for multistage stochastic linear programs with interstage dependency. Mathematical Programming **75**(2), 241–256 (1996). DOI 10.1007/BF02592154
32. Jin, S., Botterud, A., Ryan, S.M.: Temporal versus stochastic granularity in thermal generation capacity planning with wind power. IEEE Transactions on Power Systems **29**(5), 2033–2041 (2014). DOI 10.1109/TPWRS.2014.2299760
33. Jin, S., Ryan, S.M., Watson, J.P., Woodruff, D.L.: Modeling and solving a large-scale generation expansion planning problem under uncertainty. Energy Systems **2**(3), 209–242 (2011). DOI 10.1007/s12667-011-0042-9
34. Kaut, M., Midthun, K.T., Werner, A.S., Tomasgard, A., Hellemo, L., Fodstad, M.: Multi-horizon stochastic programming. Computational Management Science **11**(1), 179–193 (2014). DOI 10.1007/s10287-013-0182-6
35. Kim, K., Zavala, V.M.: Algorithmic innovations and software for the dual decomposition method applied to stochastic mixed-integer programs. Mathematical Programming Computation **10**(2), 225–266 (2018). DOI 10.1007/s12532-017-0128-z
36. Koltsaklis, N., Georgiadis, M.: A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. Applied Energy **158**, 310–331 (2015). DOI 10.1016/j.apenergy.2015.08.054
37. Koltsaklis, N.E., Dagoumas, A.S.: State-of-the-art generation expansion planning: A review. Applied Energy **230**, 563 – 589 (2018). DOI https://doi.org/10.1016/j.apenergy.2018.08.087
38. Lara, C.L.: Sddip implementation for a generation expansion planning model. https://github.com/cristianallara/SDDiP (2019)
39. Lara, C.L., Mallapragada, D.S., Papageorgiou, D.J., Venkatesh, A., Grossmann, I.E.: Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. European Journal of Operational Research **271**(3), 1037 – 1054 (2018). DOI https://doi.org/10.1016/j.ejor.2018.05.039
40. Lara, C.L., Omell, B., Miller, D., Grossmann, I.E.: Expanding the scope of electric power infrastructure planning. In: M.R. Eden, M.G. Ierapetritou, G.P. Towler (eds.) 13th International Symposium on Process Systems Engineering (PSE 2018), *Computer Aided Chemical Engineering*, vol. 44, pp. 1309 – 1314. Elsevier (2018). DOI https://doi.org/10.1016/B978-0-444-64241-7.50213-5
41. Lassner, C.: pymp. https://github.com/classner/pymp (2018)
42. Li, S.: Robust optimization of electric power generation expansion planning considering uncertainty of climate change. Ph.D. thesis, Rutgers, The State University of New Jersey (2014)
43. Li, S., Coit, D., Selcuklu, S., Felder, F.: Electric power generation expansion planning: Robust optimization considering climate change. In: IIE Annual Conference and Expo 2014, pp. 1049–1058 (2014)
44. Liu, Y., Sioshansi, R., Conejo, A.J.: Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. IEEE Transactions on Power Systems **33**(1), 781–791 (2018). DOI 10.1109/TPWRS.2017.2694612
45. Lohmann, T., Hering, A.S., Rebennack, S.: Spatio-temporal hydro forecasting of multireservoir inflows for hydro-thermal scheduling. European Journal of Operational Research **255**(1), 243 – 258 (2016). DOI https://doi.org/10.1016/j.ejor.2016.05.011
46. Lohmann, T., Rebennack, S.: Tailored benders decomposition for a long-term power expansion model with short-term demand response. Management Science **63**(6), 2027–2048 (2017). DOI 10.1287/mnsc.2015.2420. URL https://doi.org/10.1287/mnsc.2015.2420
47. Lopez, J.A., Ponnambalam, K., Quintana, V.H.: Generation and transmission expansion under risk using stochastic programming. IEEE Transactions on Power Systems **22**(3), 1369–1378 (2007). DOI 10.1109/TPWRS.2007.901741
48. Loulou, R., Goldstein, G., Noble, K.: Documentation for the MARKAL Family of Models. Tech. rep., International Energy Agency (2004)
49. Loulou, R., Remne, U., Kanudia, A., Lehtila, A., Goldstein, G.: Documentation for the TIMES Model - PART I. Tech. rep., International Energy Agency (2005)
50. Luo, X., Wang, J., Dooner, M., Clarke, J.: Overview of current development in electrical energy storage technologies and the application potential in power system operation. Applied Energy **137**, 511 – 536 (2015). DOI https://doi.org/10.1016/j.apenergy.2014.09.081

51. Malcolm, S.A., Zenios, S.A.: Robust optimization for power systems capacity expansion under uncertainty. The Journal of the Operational Research Society **45**(9), 1040–1049 (1994). URL http://www.jstor.org/stable/2584145
52. Mallapragada, D.S., Papageorgiou, D.J., Venkatesh, A., Lara, C.L., Grossmann, I.E.: Impact of model resolution on scenario outcomes for electricity sector system expansion. Energy **163**, 1231 – 1244 (2018). DOI https://doi.org/10.1016/j.energy.2018.08.015
53. Mejia-Giraldo, D.: Robust and flexible planning of power system generation capacity. Ph.D. thesis, Iowa State University (2013)
54. Meja-Giraldo, D., McCalley, J.D.: Maximizing future flexibility in electric generation portfolios. IEEE Transactions on Power Systems **29**(1), 279–288 (2014). DOI 10.1109/TPWRS.2013.2280840
55. Homem-de Mello, T., de Matos, V.L., Finardi, E.C.: Sampling strategies and stopping criteria for stochastic dual dynamic programming: a case study in long-term hydrothermal scheduling. Energy Systems **2**(1), 1–31 (2011). DOI 10.1007/s12667-011-0024-y
56. Moreira, A., Pozo, D., Street, A., Sauma, E.: Reliable renewable generation and transmission expansion planning: Co-optimizing system's resources for meeting renewable targets. IEEE Transactions on Power Systems **32**(4), 3246–3257 (2017). DOI 10.1109/TPWRS.2016.2631450
57. Mulvey, J.M., Ruszczyski, A.: A new scenario decomposition method for large-scale stochastic optimization. Operations Research **43**(3), 477–490 (1995). DOI 10.1287/opre.43.3.477
58. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust optimization of large-scale systems. Operations Research **43**(2), 264–281 (1995). DOI 10.1287/opre.43.2.264
59. Munoz, F.D., Watson, J.P.: A scalable solution framework for stochastic transmission and generation planning problems. Computational Management Science **12**(4), 491–518 (2015). DOI 10.1007/s10287-015-0229-y
60. O'Neill, R.P., Krall, E.A., Hedman, K.W., Oren, S.S.: A model and approach to the challenge posed by optimal power systems planning. Mathematical Programming **140**(2), 239–266 (2013). DOI 10.1007/s10107-013-0695-3
61. Oree, V., Sayed Hassen, S., Fleming, P.: Generation expansion planning optimisation with renewable energy integration: A review. Renewable and Sustainable Energy Reviews **69**, 790–803 (2017). DOI 10.1016/j.rser.2016.11.120
62. Palmintier, B., Webster, M.: Impact of unit commitment constraints on generation expansion planning with renewables. In: 2011 IEEE Power and Energy Society General Meeting, pp. 1–7 (2011). DOI 10.1109/PES.2011.6038963
63. Palmintier, B., Webster, M.: Heterogeneous unit clustering for efficient operational flexibility modeling. In: 2014 IEEE PES General Meeting — Conference Exposition, pp. 1–1 (2014). DOI 10.1109/PESGM.2014.6939001
64. Park, H., Baldick, R.: Multi-year stochastic generation capacity expansion planning under environmental energy policy. Applied Energy **183**, 737 – 745 (2016). DOI https://doi.org/10.1016/j.apenergy.2016.08.164. URL http://www.sciencedirect.com/science/article/pii/S0306261916312739
65. Pereira, M.V.F., Pinto, L.M.V.G.: Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach. Water Resources Research **21**(6), 779–792 (1985). DOI 10.1029/WR021i006p00779. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR021i006p00779
66. Pereira, M.V.F., Pinto, L.M.V.G.: Multi-stage stochastic optimization applied to energy planning. Mathematical Programming **52**(1), 359–375 (1991). DOI 10.1007/BF01582895
67. Philpott, A., Guan, Z.: On the convergence of stochastic dual dynamic programming and related methods. Operations Research Letters **36**(4), 450 – 455 (2008). DOI https://doi.org/10.1016/j.orl.2008.01.013
68. Pina, A., Silva, C.A., Ferro, P.: High-resolution modeling framework for planning electricity systems with high penetration of renewables. Applied Energy **112**, 215 – 223 (2013). DOI http://dx.doi.org/10.1016/j.apenergy.2013.05.074. URL http://www.sciencedirect.com/science/article/pii/S030626191300487X
69. Pinto, R.J., Borges, C.T., Maceira, M.E.P.: An efficient parallel algorithm for large scale hydrothermal system operation planning. IEEE Transactions on Power Systems **28**(4), 4888–4896 (2013). DOI 10.1109/TPWRS.2012.2236654
70. Poncelet, K., Delarue, E., Duerinck, J., Six, D., D'haeseleer, W.: The importance of integrating the variability of renewables in long-term energy planning models. KU Leuven pp. 1–18 (2014)
71. Pozo, D., Contreras, J., Sauma, E.E.: Unit commitment with ideal and generic energy storage units. IEEE Transactions on Power Systems **29**(6), 2974–2984 (2014). DOI 10.1109/TPWRS.2014.2313513

72. de Queiroz, A.R., Morton, D.P.: Sharing cuts under aggregated forecasts when decomposing multi-stage stochastic programs. Operations Research Letters **41**(3), 311 – 316 (2013). DOI https://doi.org/10.1016/j.orl.2013.03.003
73. Rachev, S.T., Roemisch, W.: Quantitative stability in stochastic programming: The method of probability metrics. Mathematics of Operations Research **27**(4), 792–818 (2002). URL http://www.jstor.org/stable/3690468
74. Rebennack, S.: Generation expansion planning under uncertainty with emissions quotas. Electric Power Systems Research **114**, 78 – 85 (2014). DOI https://doi.org/10.1016/j.epsr.2014.04.010
75. Rebennack, S.: Combining sampling-based and scenario-based nested benders decomposition methods: Application to stochastic dual dynamic programming. Math. Program. **156**(1-2), 343–389 (2016). DOI 10.1007/s10107-015-0884-3
76. Rebennack, S., Flach, B., Pereira, M.V.F., Pardalos, P.M.: Stochastic hydro-thermal scheduling underco_2emissions constraints. IEEE Transactions on Power Systems **27**(1), 58–68 (2012). DOI 10.1109/TPWRS.2011.2140342
77. Sadeghi, H., Rashidinejad, M., Abdollahi, A.: A comprehensive sequential review study through the generation expansion planning. Renewable and Sustainable Energy Reviews **67**, 1369–1394 (2017). DOI 10.1016/j.rser.2016.09.046
78. Schmidt, O., Hawkes, A., Gambhir, A., Staffell, I.: The future cost of electrical energy storage based on experience rates. Nature Energy **6**, 17110 (2017). DOI 10.1038/nenergy.2017.110
79. Shapiro, A., Tekaya, W., Costa, J.P.d., Soares, M.P.: Risk neutral and risk averse stochastic dual dynamic programming method. European Journal of Operational Research **224**(2), 375 – 391 (2013). DOI https://doi.org/10.1016/j.ejor.2012.08.022
80. Short, W., Sullivan, P., Mai, T., Mowers, M., Uriarte, C., Blair, N., Heimiller, D., Martinez, A.: Regional Energy Deployment System (ReEDS). Tech. Rep. December, National Renewable Energy Laboratory (NREL) (2011). DOI NREL/TP-6A20-46534
81. Shortt, A., O'Malley, M.: Impact of variable generation in generation resource planning models. In: IEEE PES General Meeting, pp. 1–6 (2010). DOI 10.1109/PES.2010.5589461
82. Thome, F., Pereira, M., Granville, S., Fampa, M.: Non-convexities representation on hydrothermal operation planning using sddp. Unpublished pp. 1–9 (2013)
83. Union of Concerned Scientists: The Nuclear Power Dilemma. Tech. rep., Union of Concerned Scientists (2018). URL https://www.ucsusa.org/sites/default/files/attach/2018/11/Nuclear-Power-Dilemma-executive-summary.pdf
84. Watson, J.P., Woodruff, D.L.: Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. Computational Management Science **8**(4), 355–370 (2011). DOI 10.1007/s10287-010-0125-4. URL https://doi.org/10.1007/s10287-010-0125-4
85. Wogrin, S., Centeno, E., Barquin, J.: Generation capacity expansion in liberalized electricity markets: A stochastic mpec approach. IEEE Transactions on Power Systems **26**(4), 2526–2532 (2011). DOI 10.1109/TPWRS.2011.2138728
86. Zhan, Y., Zheng, Q.P., Wang, J., Pinson, P.: Generation expansion planning with large amounts of wind power via decision-dependent stochastic programming. IEEE Transactions on Power Systems **32**(4), 3015–3026 (2017). DOI 10.1109/TPWRS.2016.2626958
87. Zhang, Q., Grossmann, I.E., Lima, R.M.: On the relation between flexibility analysis and robust optimization for linear systems. AIChE Journal **62**(9), 3109–3123 (2016). DOI 10.1002/aic.15221
88. Zou, J., Ahmed, S., Sun, X.A.: Partially adaptive stochastic optimization for electric power generation expansion planning. INFORMS Journal on Computing **30**(2), 388–401 (2018). DOI 10.1287/ijoc.2017.0782
89. Zou, J., Ahmed, S., Sun, X.A.: Stochastic dual dynamic integer programming. Mathematical Programming (2018). DOI 10.1007/s10107-018-1249-5

Appendix A Detailed MSIP Formulation

The detailed Multistage Stochastic Integer Programming formulation is presented below by equations (13)-(47). As mentioned before, this is an extension of the MILP model by Lara et al. [39], now including uncertain parameters:

- for the operational uncertainty we have load demand $L_{r,t,d,s,n}$ and renewable capacity factor $Cf_{i,r,t,d,s,n}$ and uncertain parameters drawn from the 2 operational profiles;
- for the strategic uncertainty we have the fuel price $P^{\text{fuel}}_{i,t,n}$ and the carbon tax $Tx^{\text{CO}_2}_{t,n}$, which are considered separately i.e. when fuel price is assumed to be uncertain then carbon tax is assumed to be deterministic, and vice versa.

Note that if an index appears in a summation or next to a $\forall$ symbol without a corresponding set, all elements in that set are assumed.

A.1 Energy balance

Constraint (13) ensures that in each sub-period s of representative day d in year t of node n, the sum of instantaneous power $p_{i,r,t,d,s,n}$ generated by generator clusters i in region r plus the difference between the power flow going from regions r' to region r, $p^{\text{flow}}_{r',r,t,d,s,n}$, and the power flowing from region r to regions r', $p^{\text{flow}}_{r,r',t,d,s,n}$, plus the power discharged from all the storage clusters j in region r, $p^{\text{discharge}}_{j,r,t,d,s,n}$, equals the load demand $L_{r,t,d,s,n}$ at that region r, plus the power being charged to the storage clusters j in region r, $p^{\text{charge}}_{j,r,t,d,s,n}$, plus a slack for curtailment of renewable generation $cu_{r,t,d,s,n}$. The distance between regions $D_{r,r'}$ assumes the midpoint for each region, and the transmission loss $T^{\text{loss}}_{r,r'}$ is approximated by a fraction loss per mile.

$$\begin{aligned}&\sum_i \left(p_{i,r,t,d,s,n}\right) + \sum_{r'\neq r}\left(p^{\text{flow}}_{r',r,t,d,s,n}\cdot\left(1-T^{\text{loss}}_{r,r'}\cdot D_{r,r'}\right) - p^{flow}_{r,r',t,d,s,n}\right) + \sum_j p^{\text{discharge}}_{j,r,t,d,s,n} \\ &= L_{r,t,d,s,n} + \sum_j p^{\text{charge}}_{j,r,t,d,s,n} + cu_{r,t,d,s,n} \qquad \forall\, r,t \in T_n, n, d, s\end{aligned} \tag{13}$$

A.2 Capacity factor

Constraint (14) limits the power outlet $p_{i,r,t,d,s,n}$ of renewable generators to be equal to a fraction $Cf_{i,r,t,d,s,n}$ of the nameplate capacity $Qg^{\text{np}}_{i,r}$ in each sub-period s of representative day d in year t of node n, where $ngo^{\text{rn}}_{i,r,t,n}$ represents the number of renewable generators that are operational in year t of node n. Due to the flexibility in sizes for renewable generators, $ngo^{\text{rn}}_{i,r,t,n}$ is relaxed to be continuous.

$$p_{i,r,t,d,s,n} = Qg^{\text{np}}_{i,r}\cdot Cf_{i,r,t,d,s,n}\cdot ngo^{\text{rn}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}^{\text{RN}}_r, r,t \in T_n, n, d, s \tag{14}$$

A.3 Unit commitment

Constraint (15) computes the number of generators that are ON, $u_{i,r,t,d,s,n}$, or in startup, $su_{i,r,t,d,s,n}$, and shutdown, $sd_{i,r,t,d,s,n}$, modes in cluster i in sub-period s of representative day d of year t of node n, and treated as integer variables.

$$u_{i,r,t,d,s,n} = u_{i,r,t,d,s-1,n} + su_{i,r,t,d,s,n} - sd_{i,r,t,d,s,n} \quad \forall\, i \in \mathscr{I}^{\text{TH}}_r, r,t \in T_n, n, d, s \tag{15}$$

A.4 Ramping limits

Constraints (16)-(17) capture the limitation on how fast thermal units can adjust their output power, $p_{i,r,t,d,s,n}$, where $Ru_i^{\max}$ is the maximum ramp-up rate, $Rd_i^{\max}$ is the maximum ramp-down rate, and $Pg_i^{\min}$ is the minimum operating limit [63].

$$\begin{aligned} p_{i,r,t,d,s,n} - p_{i,r,t,d,s-1,n} \leq \quad & Ru_i^{\max} \cdot Hs \cdot Qg_{i,r}^{\text{np}} \cdot (u_{i,r,t,d,s,n} - su_{i,r,t,d,s,n}) \\ & + \max\left(Pg_i^{\min}, Ru_i^{\max} \cdot Hs\right) \cdot Qg_{i,r}^{\text{np}} \cdot su_{i,r,t,d,s,n} \qquad (16) \\ & \forall\, i \in \mathscr{I}_r^{\text{TH}}, r,t \in T_n, n, d, s \\ p_{i,r,t,d,s-1,n} - p_{i,r,t,d,s,n} \leq \quad & Rd_i^{\max} \cdot Hs \cdot Qg_{i,r}^{\text{np}} \cdot (u_{i,r,t,d,s,n} - su_{i,r,t,d,s,n}) \\ & + \max\left(Pg_i^{\min}, Rd_i^{\max} \cdot Hs\right) \cdot Qg_{i,r}^{\text{np}} \cdot sd_{i,r,t,d,s,n} \qquad (17) \\ & \forall\, i \in \mathscr{I}_r^{\text{TH}}, r,t \in T_n, n, d, s \end{aligned}$$

A.5 Operating limits

Constraints (18)-(19) specify that each thermal generator is either OFF and outputting zero power, or ON and running within the operating limits $Pg_i^{\min} \cdot Qg_{i,r}^{\text{np}}$ and $Qg_{i,r}^{\text{np}}$. The variable $u_{i,r,t,d,s,n}$ (integer variable) represents the number of generators that are ON in cluster $i \in \mathscr{I}_r^{\text{TH}}$ at the time period t of node n, representative day d, and sub-period s. Note that onstraint (19) is modified in order to capture the need for generators to run below the maximum considering operating reserves, where $q_{i,r,t,d,s}^{\text{spin}}$ is a variable representing the spinning reserve capacity.

$$u_{i,r,t,d,s,n} \cdot Pg_i^{\min} \cdot Qg_{i,r}^{\text{np}} \leq p_{i,r,t,d,s,n} \qquad \forall\, i \in \mathscr{I}_r^{\text{TH}}, r,t \in T_n, n, d, s \qquad (18)$$

$$p_{i,r,t,d,s,n} + q_{i,r,t,d,s,n}^{\text{spin}} \leq u_{i,r,t,d,s,n} \cdot Qg_{i,r}^{\text{np}} \qquad \forall\, i \in \mathscr{I}_r^{\text{TH}}, r,t \in T_n, n, d, s \qquad (19)$$

A.6 Total operating reserve

Constraint (20) dictates that the total spinning reserve, $q_{i,r,t,d,s,n}^{\text{spin}}$, plus quick-start reserve, $q_{i,r,t,d,s,n}^{\text{Qstart}}$, must exceed the minimum operating reserve, $Op^{\min}$, which is a percentage of the load $L_{r,t,d,s,n}$ in a reserve sharing region r at each sub-period s.

$$\sum_{i \in \mathscr{I}_r^{\text{TH}}} \left(q_{i,r,t,d,s,n}^{\text{spin}} + q_{i,r,t,d,s,n}^{\text{Qstart}}\right) \geq Op^{\min} \cdot L_{r,t,d,s,n} \qquad \forall\, r,t \in T_n, n, d, s \qquad (20)$$

A.7 Total spinning reserve

Constraint (21) specifies that the total spinning reserve $q_{i,r,t,d,s,n}^{\text{spin}}$ must exceed the minimum spinning reserve, $Spin^{\min}$, which is a percentage of the load $L_{r,t,d,s,n}$ in a reserve

sharing region r at each sub-period s.

$$\sum_{i \in \mathscr{I}_r^{\mathrm{TH}}} q^{\mathrm{spin}}_{i,r,t,d,s,n} \geq Spin^{\min} \cdot L_{r,t,d,s,n} \qquad \forall\, r,t \in T_n, n, d, s \tag{21}$$

A.8 Maximum spinning reserve

Constraint (22) states that the maximum fraction of capacity of each generator cluster that can contribute to spinning reserves is given by $Frac_i^{\mathrm{spin}}$, which is a fraction of the nameplate capacity $Qg^{\mathrm{np}}_{i,r}$.

$$q^{\mathrm{spin}}_{i,r,t,d,s,n} \leq u_{i,r,t,d,s,n} \cdot Qg^{\mathrm{np}}_{i,r} \cdot Frac_i^{\mathrm{spin}} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{TH}}, r,t \in T_n, n, d, s \tag{22}$$

A.9 Maximum quick-start reserve

Constraint (23) dictates that the maximum fraction of the capacity of each generator cluster that can contribute to quick-start reserves is given by $Frac_i^{\mathrm{Qstart}}$ (fraction of the nameplate capacity $Qg^{\mathrm{np}}_{i,r}$), and that quick-start reserves can only be provided by the generators that are OFF, i.e., not active. Here the integer variable $ngo^{\mathrm{th}}_{i,r,t,n}$ represents the number of thermal generators that are operational (i.e., installed and ready to operate) at year t of node n.

$$q^{\mathrm{Qstart}}_{i,r,t,d,s,n} \leq (ngo^{\mathrm{th}}_{i,r,t,n} - u_{i,r,t,d,s,n}) \cdot Qg^{\mathrm{np}}_{i,r} \cdot Frac_i^{\mathrm{Qstart}} \quad \forall\, i \in \mathscr{I}_r^{\mathrm{TH}}, r,t \in T_n, n, d, s \tag{23}$$

A.10 Planning reserve requirement

Constraint (24) ensures that the operating capacity is greater than or equal to the annual peak load $L_t^{\max}$, plus a predefined fraction of reserve margin $R_t^{\min}$ of the annual peak load $L_t^{\max}$. Due to the due to the renewables inability to control dispatch and the uncertainty of the output, only a fraction of their nameplate capacity, referred to as the capacity value Q_i^{v} counts towards the planning reserve requirement.

$$\begin{aligned} &\sum_{i \in \mathscr{I}_r^{\mathrm{RN}}} \sum_r \left(Qg^{\mathrm{np}}_{i,r} \cdot Q_i^{\mathrm{v}} \cdot ngo^{\mathrm{rn}}_{i,r,t,n} \right) + \sum_{i \in \mathscr{I}_r^{\mathrm{TH}}} \sum_r \left(Qg^{\mathrm{np}}_{i,r} \cdot ngo^{\mathrm{th}}_{i,r,t,n} \right) \\ &\geq (1 + R_t^{\min}) \cdot L_t^{\max} \qquad \forall\, t \in T_n, n \end{aligned} \tag{24}$$

A.11 Minimum annual renewable generation requirement

Constraint (25) ensures that, in case of policy mandates, the renewable generation quota target, $RN_t^{\min}$, which is a fraction of the energy demand $ED_{t,n}$, is satisfied. If

not, i.e, if there is a deficit $def^{\mathrm{rn}}_{t,n}$ from the quota, this is subjected to a penalty that is included later in the objective function.

$$\sum_d \sum_s \left[W_d \cdot Hs \cdot \left(\sum_{i \in \mathscr{I}_r^{\mathrm{RN}}} \sum_r p_{i,r,t,d,s,n} - cu_{r,t,d,s,n} \right) \right] + def^{\mathrm{rn}}_{,n} t \geq RN^{\min}_t \cdot ED_t \qquad \forall\, t \in T_n, n \tag{25}$$

Here W_d represents the weight of the representative day d, Hs is the length of the sub-period, $cu_{r,t,d,s,n}$ is the curtailment of renewable generation, and $ED_{t,n}$ represent the energy demand in year t of node n:

$$ED_{t,n} = \sum_r \sum_d \sum_s \left(W_d \cdot Hs \cdot L_{r,t,d,s,n} \right)$$

A.12 Maximum yearly installation

Constraints (26)-(27) limit the yearly installation per generation type in each region r to an upper bound $Q^{\mathrm{inst,UB}}_{i,t}$ in MW/year. Here $ngb^{\mathrm{rn}}_{i,r,t,n}$ and $ngb^{\mathrm{th}}_{i,r,t,n}$ represent the number of renewable and thermal generators built in region r in year t of node n, respectively. Note that due to the flexibility in sizes for renewable generators, $ngb^{\mathrm{rn}}_{i,r,t,n}$ is relaxed to be continuous.

$$\sum_r ngb^{\mathrm{rn}}_{i,r,t,n} \leq Q^{\mathrm{inst,UB}}_{i,t} / Qg^{\mathrm{np}}_{i,r} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{Rnew}}, t \in T_n, n \tag{26}$$

$$\sum_r ngb^{\mathrm{th}}_{i,r,t,n} \leq Q^{\mathrm{inst,UB}}_{i,t} / Qg^{\mathrm{np}}_{i,r} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{Tnew}}, t \in T_n, n \tag{27}$$

A.13 Balance of generators

Concerning renewable generator clusters, we define a set of constraints (28)-(29) to compute the number of generators in cluster i that are ready to operate $ngo^{\mathrm{rn}}_{i,r,t,n}$, taking into account the generators that were already existing at the beginning of the planning horizon $Ng^{\mathrm{Rold}}_{i,r}$, the generators built $ngb^{\mathrm{rn}}_{i,r,t,n}$, and the generators retired $ngr^{\mathrm{rn}}_{i,r,t,n}$ at year t of node n. It is important to highlight that we assume *no lead time* between the decision to build/install a generator and the moment it can begin producing electricity.

$$ngo^{\mathrm{rn}}_{i,r,t,n} = Ng^{\mathrm{Rold}}_{i,r} + ngb^{\mathrm{rn}}_{i,r,t,n} - ngr^{\mathrm{rn}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{RN}}, r, t = 1, n = 1 \tag{28}$$

$$ngo^{\mathrm{rn}}_{i,r,t,n} = ngo^{\mathrm{rn}}_{i,r,t-1,P(n)} + ngb^{\mathrm{rn}}_{i,r,t,n} - ngr^{\mathrm{rn}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{RN}}, r, t \in T_n, n > 1 \tag{29}$$

As aforementioned, due to the flexibility in sizes for renewable generators, $ngo^{\mathrm{rn}}_{i,r,t,n}$, $ngb^{\mathrm{rn}}_{i,r,t,n}$, and $ngr^{\mathrm{rn}}_{i,r,t,n}$ are relaxed to be continuous. Note that $ngb^{\mathrm{rn}}_{i,r,t,n}$ for $i \in \mathscr{I}_r^{\mathrm{Rold}}$ is fixed to zero in all time periods, i.e., the clusters of existing renewable generators cannot have any new additions during the time horizon considered.

We also define constraint (30) to enforce the renewable generators that reached the end of their lifetime to either retire, $ngr^{\mathrm{rn}}_{i,r,t,n}$, or have their life extended, $nge^{\mathrm{rn}}_{i,r,t,n}$.

$Ng^{\mathrm{r}}_{i,r,t}$ is a parameter that represents the number of old generators (i.e., $i \in \mathscr{I}_r^{\mathrm{old}}$) that reached the end of their lifetime, LT_i, at year t.

$$Ng^{\mathrm{r}}_{i,r,t} = ngr^{\mathrm{rn}}_{i,r,t,n} + nge^{\mathrm{rn}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{Rold}}, r,t \in T_n, n \tag{30}$$

Concerning thermal generator clusters, we define a set of constraints (31)-(32) to compute the number of generators in cluster i that are ready to operate $ngo^{\mathrm{th}}_{i,r,t,n}$, taking into account the generators that were already existing at the beginning of the planning horizon $Ng^{\mathrm{Told}}_{i,r}$, the generators built $ngb^{\mathrm{th}}_{i,r,t,n}$, and the generators retired $ngr^{\mathrm{th}}_{i,r,t,n}$ at year t of node n.

$$ngo^{\mathrm{th}}_{i,r,t,n} = Ng^{\mathrm{Told}}_{i,r} + ngb^{\mathrm{th}}_{i,r,t,n} - ngr^{\mathrm{th}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{TH}}, r,t = 1, n = 1 \tag{31}$$

$$ngo^{\mathrm{th}}_{i,r,t,n} = ngo^{\mathrm{th}}_{i,r,t-1,P(n)} + ngb^{\mathrm{th}}_{i,r,t,n} - ngr^{\mathrm{th}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{TH}}, r,t \in T_n, n > 1 \tag{32}$$

Note that $ngb^{\mathrm{th}}_{i,r,t,n}$ for $i \in \mathscr{I}_r^{\mathrm{Told}}$ is fixed to zero in all time periods, i.e., the clusters of existing thermal generators cannot have any new additions during the time horizon considered.

We also define constraint (33) to enforce the thermal generators that reached the end of their lifetime to either retire, $ngr^{\mathrm{th}}_{i,r,t,n}$, or have their life extended $nge^{\mathrm{th}}_{i,r,t,n}$.

$$Ng^{\mathrm{r}}_{i,r,t} = ngr^{\mathrm{th}}_{i,r,t,n} + nge^{\mathrm{th}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{Told}}, r,t \in T_n, n \tag{33}$$

Finally, we have constraint (34) that ensures that only installed generators can be in operation:

$$u_{i,r,t,d,s,n} \leq ngo^{\mathrm{th}}_{i,r,t,n} \qquad \forall\, i \in \mathscr{I}_r^{\mathrm{Tnew}}, r,t \in T_n, n, d, s \tag{34}$$

A.14 Storage

The energy storage devices are assumed to be ideal and generic [71]. Constraints (35)-(36) compute the number of storage units that are ready to operate $nso_{j,r,t,n}$, taking into account the storage units already existing at the beginning of the planning horizon $Ns_{j,r}$ and the ones built $nsb_{j,r,t,n}$ at year t of node n. Due to the flexibility in sizes for storage units, $nso_{j,r,t,n}$ and $nsb_{j,r,t,n}$ are relaxed to be continuous.

$$nso_{j,r,t,n} = Ns_{j,r} + nsb_{j,r,t,n} \qquad \forall\, j,r,t = 1, n = 1 \tag{35}$$

$$nso_{j,r,t,n} = nso_{j,r,t-1,P(n)} + nsb_{j,r,t,n} \qquad \forall\, j,r,t \in T_n, n > 1 \tag{36}$$

Constraints (37) and (38) establish that the power charge, $p^{\mathrm{charge}}_{j,r,t,d,s,n}$, and discharge, $p^{\mathrm{discharge}}_{j,r,t,d,s,n}$, of the storage units in cluster j, $nso_{j,r,t,n}$, has to be within the operating limits: $Charge^{\mathrm{min}}_j$ and $Charge^{\mathrm{max}}_j$, and $Discharge^{\mathrm{min}}_j$ and $Discharge^{\mathrm{min}}_j$, respectively.

$$Charge^{\mathrm{min}}_j \cdot nso_{j,r,t,n} \leq p^{\mathrm{charge}}_{j,r,t,d,s,n} \leq Charge^{\mathrm{max}}_j \cdot nso_{j,r,t,n} \qquad \forall\, j,r,t \in T_n, n, d, s \tag{37}$$

$$Discharge^{\mathrm{min}}_j \cdot nso_{j,r,t,n} \leq p^{\mathrm{discharge}}_{j,r,t,d,s,n} \leq Discharge^{\mathrm{max}}_j \cdot nso_{j,r,t,n} \qquad \forall\, j,r,t \in T_n, n, d, s \tag{38}$$

Constraint (39) specifies that the energy storage level, $p^{\text{level}}_{j,r,t,d,s,n}$, for the storage units in cluster j, $nso_{j,r,t,n}$ has to be within the storage capacity limits $Storage^{\min}_{j}$ and $Storage^{\max}_{j}$.

$$Storage^{\min}_{j} \cdot nso_{j,r,t,n} \leq p^{\text{level}}_{j,r,t,d,s} \leq Storage^{\max}_{j} \cdot nso_{j,r,t,n} \quad \forall\ j,r,t \in T_n,n,d,s \quad (39)$$

Constraints (40) and (41) show the power balance in the storage units. The state of charge $p^{\text{level}}_{j,r,t,d,s,n}$ at the end of sub-period s depends on the previous state of charge $p^{\text{level}}_{j,r,t,d,s-1,n}$, and the power charged $p^{\text{charge}}_{j,r,t,d,s,n}$ and discharged $p^{\text{discharge}}_{j,r,t,d,s,n}$ at sub-period s. The symbols η^{charge}_{j} and $\eta^{\text{discharge}}_{j}$ represent the charging and discharging efficiencies, respectively. For the first hour of the day d in year t of node n, the previous state of charge (i.e., $s = 0$) is the variable $p^{\text{level,0}}_{j,r,t,d,n}$.

$$p^{\text{level}}_{j,r,t,d,s,n} = p^{\text{level}}_{j,r,t,d,s-1,n} + \eta^{\text{charge}}_{j} \cdot p^{\text{charge}}_{j,r,t,d,s} + p^{\text{discharge}}_{j,r,t,d,s,n} / \eta^{\text{discharge}}_{j} \quad \forall\ j,r,t \in T_n,n,d,s > 1 \quad (40)$$

$$p^{\text{level}}_{j,r,t,d,s,n} = p^{\text{level,0}}_{j,r,t,d,n} + \eta^{\text{charge}}_{j} \cdot p^{\text{charge}}_{j,r,t,d,s,n} + p^{\text{discharge}}_{j,r,t,d,s,n} / \eta^{\text{discharge}}_{j} \quad \forall\ j,r,t \in T_n,n,d,s = 1 \quad (41)$$

Constraints (42) and (43) force the storage units to begin $p^{\text{level,0}}_{j,r,t,d,s}$ and end $p^{\text{level}}_{j,r,t,d,s=S,n}$ each day d of year t with 50% of their maximum storage $Storage^{\max}_{j}$. This is a heuristic to attach carryover storage level form one representative day to the next [44].

$$p^{\text{level,0}}_{j,r,t,d,n} = 0.5 \cdot Storage^{\max}_{j} \cdot nso_{j,r,t,n} \quad \forall\ j,r,t \in T_n,n,d \quad (42)$$

$$p^{\text{level}}_{j,r,t,d,s,n} = 0.5 \cdot Storage^{\max}_{j} \cdot nso_{j,r,t,n} \quad \forall\ j,r,t \in T_n,n,d,s = S \quad (43)$$

A.15 Objective function

The objective of this model is to minimize the *expected* net present cost, Φ, over the planning horizon, which includes operating costs Φ^{opex}, investment costs Φ^{capex}, and potential penalties Φ^{PEN} for not meeting the the targets on renewables.

$$\min \quad \Phi = \sum_{n \in \mathscr{T}} prob_n \cdot \sum_{t \in T_n} \left(\Phi^{\text{opex}}_{t,n} + \Phi^{\text{capex}}_{t,n} + \Phi^{\text{PEN}}_{t,n} \right) \quad (44)$$

The operating expenditure, $\Phi^{\text{opex}}_{t,n}$, comprises the variable $VOC_{i,t}$ and fixed $FOC_{i,t}$ operating costs, as well as fuel cost $P^{\text{fuel}}_{i,t,n}$ per heat rate HR_i, carbon tax $Tx^{\text{CO}_2}_{t,n}$ for CO_2 emissions $EF^{CO_2}_{i}$, and start-up cost (variable cost $P^{\text{fuel}}_{i,t,n}$ that depends on the amount of fuel burned for startup F^{start}_{i}, and fixed cost C^{start}_{i}). Both $P^{\text{fuel}}_{i,t,n}$ and $Tx^{\text{CO}_2}_{t,n}$ are potential

strategic uncertain parameters, hence are indexed by node n.

$$\begin{aligned}\Phi_{t,n}^{\text{opex}} = \; & If_t \cdot \Bigg[\sum_d \sum_s W_d \cdot hs \cdot \\ & \left(\sum_i \sum_r (VOC_{i,t} + P_{i,t,n}^{\text{fuel}} \cdot HR_i + Tx_{t,n}^{\text{CO}_2} \cdot EF_i^{CO_2} \cdot HR_i) \cdot p_{i,r,t,d,s,n} \right) \\ & + \left(\sum_{i \in \mathscr{I}_r^{\text{RN}}} \sum_r FOC_{i,t} \cdot Qg_{i,r}^{\text{np}} \cdot ngo_{i,r,t,n}^{\text{rn}} \right) \\ & + \left(\sum_{i \in \mathscr{I}_r^{\text{TH}}} \sum_r FOC_{i,t} \cdot Qg_{i,r}^{\text{np}} \cdot ngo_{i,r,t,n}^{\text{th}} \right) \\ & + \sum_{i \in \mathscr{I}_r^{\text{TH}}} \sum_r \sum_d \sum_s W_d \cdot Hs \cdot su_{i,r,t,d,s,n} \cdot Qg_{i,r}^{\text{np}} \\ & \cdot \left(F_i^{\text{start}} \cdot P_{i,t,n}^{\text{fuel}} + F_i^{\text{start}} \cdot EF^{CO_2} \cdot Tx_{t,n}^{\text{CO}_2} + C_i^{\text{start}} \right) \Bigg] \end{aligned} \tag{45}$$

The capital expenditure, $\Phi_{t,n}^{\text{capex}}$, includes the amortized cost of acquiring new generators, $DIC_{i,t}$, new storage devices, $SIC_{j,t}$, and the amortized cost of extending the life of generators that reached their expected lifetime. The latter is assumed to be a fraction LE_i of the investment cost, $DIC_{i,t}$, in a new generator with the same or equivalent generation technology. In this framework, the investment cost takes into account the remaining value at the end of the time horizon by considering the annualized capital cost and multiplying it by the number of years remaining in the planning horizon at the time of installation to calculate the $DIC_{i,t}$.

$$\begin{aligned}\Phi_{t,n}^{\text{capex}} = \; & If_t \cdot \Bigg[\sum_{i \in \mathscr{I}_r^{\text{Rnew}}} \sum_r DIC_{i,t} \cdot CC_i^{\text{m}} \cdot Qg_{i,r}^{\text{np}} \cdot ngb_{i,r,t,n}^{\text{rn}} \\ & + \sum_{i \in \mathscr{I}_r^{\text{Tnew}}} \sum_r DIC_{i,t} \cdot CC_i^{\text{m}} \cdot Qg_{i,r}^{\text{np}} \cdot ngb_{i,r,t,n}^{\text{th}} \\ & + \sum_j \sum_r SIC_{j,t} \cdot Storage_j^{\max} \cdot nsb_{j,r,t,n} \\ & + \sum_{i \in \mathscr{I}_r^{\text{RN}}} \sum_r DIC_{i,t} \cdot LE_i \cdot Qg_{i,r}^{\text{np}} \cdot nge_{i,r,t,n}^{\text{rn}} \\ & + \sum_{i \in \mathscr{I}_r^{\text{TH}}} \sum_r DIC_{i,t} \cdot LE_i \cdot Qg_{i,r}^{\text{np}} \cdot nge_{i,r,t,n}^{\text{th}} \Bigg] \end{aligned} \tag{46}$$

The capital multiplier CC_i^{m} associated with new generator clusters is meant to account for differences in depreciation schedules applicable to each technology, with higher values being indicative of slower depreciating schedule and vice versa.

Lastly, the penalty cost, $\Phi^{\text{PEN}}_{t,n}$, includes the potential fines for not meeting the renewable energy quota, PEN^{rn}_t, and curtailing the renewable generation.

$$\Phi^{\text{PEN}}_{t,n} = \quad If_t \cdot \left(PEN^{\text{rn}}_t \cdot def^{\text{rn}}_{t,n} + PEN^c \cdot \sum_r \sum_d \sum_s cu_{r,t,d,s,n} \right) \tag{47}$$

Appendix B Nomenclature

B.1 Indices and Sets

$r \in R$	set of regions within the area considered
$i \in I$	set of generator clusters
$i \in I_r$	set of generator clusters in region r
$i \in I^{\text{old}}_r$	set of existing generator clusters in region r at the beginning of the time horizon, $I^{\text{old}}_r \subset I_r$
$i \in I^{\text{new}}_r$	set of potential generator clusters in region r, $I^{\text{new}}_r \subseteq I_r$
$i \in I^{\text{TH}}_r$	set of thermal generator clusters in region r, $I^{\text{TH}}_r \subset I_r$
$i \in I^{\text{RN}}_r$	set of renewable generator clusters in region r, $I^{\text{RN}}_r \subset I_r$
$i \in I^{\text{Told}}_r$	set of existing thermal generator clusters in region r, $I^{\text{Told}}_r \subset I^{\text{TH}}_r$
$i \in I^{\text{Tnew}}_r$	set of potential thermal generator clusters in region r, $I^{\text{Tnew}}_r \subset I^{\text{TH}}_r$
$i \in I^{\text{Rold}}_r$	set of existing renewable generator clusters in region r, $I^{\text{Rold}}_r \subset I^{\text{RN}}_r$
$i \in I^{\text{Rnew}}_r$	set of potential renewable generator clusters in region r, $I^{\text{Rnew}}_r \subset I^{\text{RN}}_r$
$j \in J$	set of storage unit clusters
$n \in \mathscr{T}$	set of nodes in the scenario tree $\mathscr{T}$
$t \in T_n$	set of time periods (years) within each node of the scenario tree $\mathscr{T}$
$d \in D$	set of representative days in each year t
$s \in S$	set of sub-periods of time per representative day d in year t
$k \in K$	set of iterations in the Nested Decomposition algorithm

B.2 Parameters

$L_{r,t,d,s,n}$	load demand in region r in sub-period s of representative day d of year t of node n (MW)
$L^{\max}_t$	peak load in year t (MW)
W_d	weight of the representative day d
Hs	duration of sub-period s (hours)
$Qg^{\text{np}}_{i,r}$	nameplate (nominal) capacity of a generator in cluster i in region r (MW)
$Ng^{\text{old}}_{i,r}$	number of existing generators in each cluster, $i \in \mathscr{I}^{\text{old}}_r$, per region r at the beginning of the time horizon
$Ng^{\max}_i$	maximum number of generators in the potential clusters $i \in \mathscr{I}^{\text{new}}_r$
$Q^{\text{inst,UB}}_{i,t}$	upper bound on yearly capacity installations based on generation technology ($^{\text{MW}}/_{\text{year}}$)

$R_t^{\min}$	system's minimum reserve margin for year t (fraction of the peak load)
ED_t	energy demand during year t (MWh)
LT_i	expected lifetime of generation cluster i (years)
T_t^{remain}	remaining time until the end of the time horizon at year t (years)
$Ng_{i,r,t}^{\text{r}}$	number of generators in cluster i of region r that achieved their expected lifetime
Q_i^{v}	capacity value of generation cluster i (fraction of the nameplate capacity)
$Cf_{i,r,t,d,s,n}$	capacity factor of generation cluster $i \in \mathscr{I}_r^{\text{RN}}$ in region r at sub-period s, of representative day d of year t of node n (fraction of the nameplate capacity)
$Pg_i^{\min}$	minimum operating output of a generator in cluster $i \in \mathscr{I}_r^{\text{TH}}$ (fraction of the nameplate capacity)
$Ru_i^{\max}$	maximum ramp-up rate for cluster $i \in \mathscr{I}_r^{\text{TH}}$ (fraction of nameplate capacity)
$Rd_i^{\max}$	maximum ramp-down rate for cluster $i \in \mathscr{I}_r^{\text{TH}}$ (fraction of nameplate capacity)
F_i^{start}	fuel usage at startup (MMbtu/MW)
$Frac_i^{\text{spin}}$	maximum fraction of nameplate capacity of each generator that can contribute to spinning reserves (fraction of nameplate capacity)
$Frac_i^{\text{Qstart}}$	maximum fraction of nameplate capacity of each generator that can contribute to quick-start reserves (fraction of nameplate capacity)
$Op^{\min}$	minimum total operating reserve (fraction of the load demand)
$Spin^{\min}$	minimum spinning operating reserve (fraction of the load demand)
$Qstart^{\min}$	minimum quick-start operating reserve (fraction of the load demand)
α^{RN}	fraction of the renewable generation output covered by quick-start reserve (fraction of total renewable power output)
$T_{r,r'}^{\text{loss}}$	transmission loss factor between region r and region $r' \neq r$ ($\%/\text{miles}$)
$D_{r,r'}$	distance between region r and region $r' \neq r$ (miles)
$Ns_{j,r}$	number of existing storage units in each cluster j per region r at the beginning of the time horizon
$Charge_j^{\min}$	minimum operating charge for storage unit in cluster j (MW)
$Charge_j^{\max}$	maximum operating charge for storage unit in cluster j (MW)
$Discharge_j^{\min}$	minimum operating discharge for storage unit in cluster j (MW)
$Discharge_j^{\max}$	maximum operating discharge for storage unit in cluster j (MW)
$Storage_j^{\min}$	minimum storage capacity for storage unit in cluster j (MWh)
$Storage_j^{\max}$	maximum storage capacity (i.e. nameplate capacity) for storage unit in cluster j (MWh)
η_j^{charge}	charging efficiency of storage unit in cluster j (fraction)
$\eta_j^{\text{discharge}}$	discharging efficiency of storage unit in cluster j (fraction)

LT^{s}_{j}	lifetime of storage unit in cluster j (years)
Ir	nominal interest rate
If_{t}	discount factor for year t
$OCC_{i,t}$	overnight capital cost of generator cluster i in year t ($/MW)
$ACC_{i,t}$	annualized capital cost of generator cluster i in year t ($/MW)
$DIC_{i,t}$	discounted investment cost of generator cluster i in year t ($/MW) [2]
$SIC_{j,t}$	investment cost of storage cluster j in year t ($/MW)
CC^{m}_{i}	capital cost multiplier of generator cluster i (unitless)
LE_{i}	life extension cost for generator cluster i (fraction of the investment cost of corresponding new generator)
$FOC_{i,t}$	fixed operating cost of generator cluster i ($/MW)
$P^{\mathrm{fuel}}_{i,t,n}$	price of fuel for generator cluster i in year t of node n ($/MMBtu)
HR_{i}	heat rate of generator cluster i (MMBtu/MWh)
$Tx^{\mathrm{CO_2}}_{t,n}$	carbon tax in year t of node n ($/kg CO_2)
$EF^{\mathrm{CO_2}}_{i}$	full lifecycle CO_2 emission factor for generator cluster i ($kgCO_2$/MMBtu)
$VOC_{i,t}$	variable O&M cost of generator cluster i ($/MWh)
$RN^{\min}_{t}$	minimum renewable energy production requirement during year t (fraction of annual energy demand)
PEN^{rn}_{t}	penalty for not meeting renewable energy quota target during year t ($/MWh)
PEN^{c}_{t}	penalty for curtailment during year t ($/MWh)
C^{start}_{i}	fixed startup cost for generator cluster i ($/MW)

B.3 Continuous variables

Φ	net present cost[3] throughout the time horizon, including amortized investment cost, operational and environmental cost ($)
$\Phi^{\mathrm{opex}}_{t,n}$	amortized operating costs in year t of node n ($)
$\Phi^{\mathrm{capex}}_{t,n}$	amortized investment costs in year t of node n ($)
$\Phi^{\mathrm{PEN}}_{t,n}$	amortized penalty costs in year t of node n ($)
$p_{i,r,t,d,s,n}$	power output of generation cluster i in region r during sub-period s of representative day d of year t of node n (MW)
$def^{\mathrm{rn}}_{t,n}$	deficit from renewable energy quota target during year t of node n (MWh)
$cu_{r,t,ss,s,n}$	curtailment slack generation in region r during sub-period s of representative day d of year t of node n (MW)
$p^{\mathrm{flow}}_{r,r',t,d,s,n}$	power transfer from region r to region $r' \neq r$ during sub-period s of representative day d of year t of node n (MW)
$q^{\mathrm{spin}}_{i,r,t,d,s,n}$	spinning reserve capacity of generation cluster i in region r during sub-period s of representative day d of year t of node n (MW)

[2] $DIC_{i,t}$ is used in the calculation for the life extension investment cost, which is in terms of a fraction LE_i of the capital cost. Therefore the investment cost for the existing cluster is approximated as being the same as for the potential clusters that have the same or similar generation technology.

[3] All the costs are in 2015 USD.

$q^{\text{Qstart}}_{i,r,t,d,s,n}$ quick-start capacity reserve of generation cluster i in region r during sub-period s of representative day d of year t of node n (MW)

$ngo^{\text{rn}}_{i,r,t,n}$ number of generators that are operational in cluster $i \in \mathscr{I}_r^{\text{RN}}$ of region r in year t of node n (continuous relaxation)

$ngb^{\text{rn}}_{i,r,t,n}$ number of generators that are built in cluster $i \in \mathscr{I}_r^{\text{RN}}$ of region r in year t of node n (continuous relaxation)

$ngr^{\text{rn}}_{i,r,t,n}$ number of generators that retire in cluster $i \in \mathscr{I}_r^{\text{RN}}$ of region r in year t of node n (continuous relaxation)

$nge^{\text{rn}}_{i,r,t,n}$ number of generators that had their life extended in cluster $i \in \mathscr{I}_r^{\text{RN}}$ of region r in year t of node n (continuous relaxation)

$p^{\text{charge}}_{j,r,t,d,s,n}$ power being charged to storage cluster j is region r, during sub-period s of representative day d of year t of node n (MW)

$p^{\text{discharge}}_{j,r,t,d,s,n}$ power being discharged to storage cluster j is region r, during sub-period s of representative day d of year t of node n (MW)

$p^{\text{level}}_{j,r,t,d,s,n}$ state of charge of storage cluster j is region r, during sub-period s of representative day d of year t of node n (MWh)

$p^{\text{level,0}}_{j,r,t,d,n}$ state of charge of storage cluster j is region r at hour zero of representative day d of year t of node n (MWh)

$nso_{j,r,t,n}$ number of storage units that are operational in cluster j of region r in year t of node n (continuous relaxation)

$nsb_{j,r,t,n}$ number of storage units that are built in cluster j of region r in year t of node n (continuous relaxation)

$nsr_{j,r,t,n}$ number of storage units that retire in cluster j of region r in year t of node n (continuous relaxation)

B.4 Discrete variables

$ngo^{\text{th}}_{i,r,t,n}$ number of generators that are operational in cluster $i \in \mathscr{I}_r^{\text{TH}}$ of region r in year t of node n (integer variable)

$ngb^{\text{th}}_{i,r,t,n}$ number of generators that are built in cluster $i \in \mathscr{I}_r^{\text{TH}}$ of region r in year t of node n (integer variable)

$ngr^{\text{th}}_{i,r,t,n}$ number of generators that retire in cluster $i \in \mathscr{I}_r^{\text{TH}}$ of region r in year t of node n (integer variable)

$nge^{\text{th}}_{i,r,t,n}$ number of generators that had their life extended in cluster $i \in \mathscr{I}_r^{\text{TH}}$ of region r in year t of node n (integer variable)

$u_{i,r,t,d,s,n}$ number of thermal generators ON in cluster $i \in \mathscr{I}_r$ of region r during sub-period s of representative day d of year t of node n (integer variable)

$su_{i,r,t,d,s,n}$ number of generators starting up in cluster i during sub-period s of representative day d in year t of node n (integer variable)

$sd_{i,r,t,d,s,n}$ number of generators shutting down in cluster i during sub-period s of representative day d in year t of node n (integer variable)