Applied Energy 291 (2021) 116736

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A distributed computing framework for multi-stage stochastic planning of renewable power systems with energy storage as flexibility option

Angela Flores-Quiroz [a,b,*], Kai Strunz [a]

[a] Chair of Sustainable Electric Networks and Sources of Energy, Technische Universität Berlin, Germany
[b] Department of Electrical Engineering, Universidad de Chile, Chile

ARTICLE INFO

Keywords:
Power system planning
Stochastic optimization
Renewable energy
Energy storage
Operational flexibility
Distributed computing

ABSTRACT

An integrated generation, transmission, and energy storage planning model accounting for short-term constraints and long-term uncertainty is proposed. The model allows to accurately quantify the value of flexibility options in renewable power systems by representing short-term operation through the unit commitment constraints. Long-term uncertainty is represented through a scenario tree. The resulting model is a large-scale multi-stage stochastic mixed-integer programming problem. To overcome the computational burden, a distributed computing framework based on the novel Column Generation and Sharing algorithm is proposed. The performance improvement of the proposed approach is demonstrated through study cases applied to the NREL 118-bus power system. The results confirm the added value of modeling short-term constraints and long-term uncertainty simultaneously. The computational case studies show that the proposed solution approach clearly outperforms the state of the art in terms of computational performance and accuracy. The proposed planning framework is used to assess the value of energy storage systems in the transition to a low-carbon power system.

1. Introduction

Low-carbon power systems of the future will be dominated by renewable energy sources (RES). A large utilization of variable RES will increase the flexibility requirements of power systems [1]. Flexibility is defined as the ability of power systems to adapt to variability and uncertainty in demand and generation [2,3]. Flexibility is a critical aspect to be considered in the capacity planning process [4]. Neglecting short-term details in long-term planning leads to an underestimation of the flexibility requirements and can result in suboptimal or even infeasible generation portfolios [5]. Neglecting short-term details also underestimates the value of flexibility options, such as energy storage systems (ESSs) [6]. Therefore, short-term flexibility issues should be represented in long-term planning models to efficiently integrate high levels of RES. Furthermore, the coordination between generation and transmission planning is considered to have a substantial impact on the optimal generation-flexibility mix [4,7]. Since investments are made for the long term, with several parameters being uncertain at the moment of decision, there is a strong need to represent long-term uncertainty to anticipate future scenarios and avoid the risk of locking into inefficient investment decisions [7].

Including short-term flexibility requirements in long-term planning requires a finer representation of the hourly operation of the power system. This leads to an increase in the complexity of the model. Therefore, different approaches have been proposed to consider short-term constraints in long-term planning while maintaining the problem computationally tractable [5,8–11]. In [8], the unit commitment (UC) problem was included in generation expansion planning. To avoid tractability issues, a reduced number of units and a planning horizon of a single year represented by five typical weeks were considered. In [9], the model complexity was reduced by considering a linear version of the UC problem and a reduced set of operating periods. Ref. [5] proposed clustering generators into equivalent plants to reduce the number of binary variables, obtaining a 99 % reduction in the solution time. In [10], complexity was overcome by using a convex relaxation of the UC problem to represent the operation of the power system. Results show that the formulation is tractable, and the solution is more accurate compared with [5]. In [11], transmission constraints were included in the clustered UC. Tractability was maintained by relaxing the integrality of the commitment state. In [12], a generation and transmission planning model including electric vehicles and ramping constraints was proposed. The model was kept tractable by considering six representative days to model the yearly operation. Although the methods presented in [5,8–12] reported clear improvements, only

* Corresponding author at: Chair of Sustainable Electric Networks and Sources of Energy, Technische Universität Berlin, Germany.
E-mail addresses: angela.flores@campus.tu-berlin.de (A. Flores-Quiroz), kai.strunz@tu-berlin.de (K. Strunz).

https://doi.org/10.1016/j.apenergy.2021.116736
Received 30 September 2020; Received in revised form 9 February 2021; Accepted 21 February 2021
Available online 22 March 2021
0306-2619/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

Investment decisions

$\boldsymbol{I}_n$	Vector of additional units installed in node n.
$I^{\mathrm{E}}_{n,e}$	Additional units of ESS e installed in node n.
$I^{\mathrm{G}}_{n,g}$	Additional units of generator g installed in node n.
$I^{\mathrm{L}}_{n,l}$	Additional lines l installed in node n.
$\boldsymbol{Z}_n$	Vector of total installed units in node n.
$Z^{\mathrm{E}}_{n,e}$	Total installed units of ESS e in node n.
$Z^{\mathrm{G}}_{n,g}$	Total installed units of generator g in node n.
$Z^{\mathrm{L}}_{n,l}$	Total installed lines l in node n.

Operational decisions

$D_{n,t,g}$	# shutdowns for generator g at hour t in node n.
$E_{n,t,e}$	Energy stored in ESS e at hour t in node n [MWh].
$F_{n,t,l}$	Flow on line l at hour t in node n [MW].
$L^{\mathrm{S}}_{n,t,b}$	Load shedding at bus b at hour t in node n [MW].
$P_{n,t,g}$	Output power of generator g at hour t in node n [MW].
$P^{\mathrm{c}}_{n,t,e}$	Power consumed by ESS e at hour t in node n [MW].
$P^{\mathrm{d}}_{n,t,e}$	Power delivered by ESS e at hour t in node n [MW].
$R^{\mathrm{prim}}_{n,t,e}$	Primary reserve of ESS e at hour t in node n [MW].
$R^{\mathrm{prim}}_{n,t,g}$	Primary reserve of generator g at hour t in node n [MW].
$\underline{R}^{\mathrm{sec}}_{n,t,e}$	Downward secondary reserve of ESS e at hour t in node n [MW].
$\underline{R}^{\mathrm{sec}}_{n,t,g}$	Downward secondary reserve of generator g at hour t in node n [MW].
$\overline{R}^{\mathrm{sec}}_{n,t,e}$	Upward secondary reserve of ESS e at hour t in node n [MW].
$\overline{R}^{\mathrm{sec}}_{n,t,g}$	Upward secondary reserve of generator g at hour t in node n [MW].
$S_{n,t,g}$	# of start-ups of generator g at hour t in node n.
$U_{n,t,g}$	# of committed units of generator g at hour t in node n.
$U^{\mathrm{Ec}}_{n,t,e}$	# of units of ESS e charging at hour t in node n.
$U^{\mathrm{Ed}}_{n,t,e}$	# of units of ESS e discharging at hour t in node n.
$\boldsymbol{X}_n$	Vector of operational variables in node n.

Parameters

$\alpha^{\mathrm{avail}}_{n,t,g}$	% of available generation of one unit of generator g at hour t in node n.
α^{R}	% of variable generation output covered by secondary reserve.
β^{Rp}	Minimum duration of primary reserve [h].
β^{Rs}	Minimum duration of secondary reserve [h].
γ_e	Emission factor of ESS e [t CO_2/MWh].
γ_g	Emission factor of generator g [t CO_2/MWh].
Γ^{max}_n	Maximum CO_2 emission in node n [t CO_2].
Δ^{Rp}	Time interval by which primary reserve must be fully active [h].
Δ^{Rs}	Time interval by which secondary reserve must be fully active [h].
η^{c}_e	Charging efficiency of ESS e.
η^{d}_e	Discharging efficiency of ESS e.
η^{sd}_e	Self discharge per hour of ESS e [h^{-1}].
τ^{off}_g	Minimum off time of generator g [h].
τ^{on}_g	Minimum on time of generator g [h].
ϕ_n	Probability of node n.
$\boldsymbol{A}_n$	Matrix coupling operational and investment decisions in node n.
$\boldsymbol{c}^{\mathrm{inv}}_n$	Vector of investment cost in node n.
$c^{\mathrm{inv}}_{n,e}$	Investment cost of ESS e in node n [EUR/MW].
$c^{\mathrm{inv}}_{n,g}$	Investment cost of generator g in node n [EUR/MW].
$c^{\mathrm{inv}}_{n,l}$	Investment cost of line l in node n [EUR/MW].
$\boldsymbol{c}^{\mathrm{op}}_n$	Vector of operational cost in node n.
c^{s}_g	Start-up cost of generator g [EUR].
c^{UD}_n	Unsupplied demand cost in node n [EUR/MWh].
$c^{\mathrm{var}}_{n,e}$	Variable cost of ESS e in node n [EUR/MWh].
$c^{\mathrm{var}}_{n,g}$	Variable cost of generator g in node n [EUR/MWh].
$d_{n,t,b}$	Load at bus b at hour t in node n [MW].
E^{max}_e	Maximum energy of a unit of ESS e [MWh].
E^{min}_e	Minimum energy of a unit of ESS e [MWh].
F^{max}_l	Maximum capacity of line l [MW].
$P^{\mathrm{c\,max}}_e$	Maximum charge power of a unit of ESS e [MW].
$P^{\mathrm{c\,min}}_e$	Minimum charge power of a unit of ESS e [MW].
$P^{\mathrm{d\,max}}_e$	Maximum discharge power of a unit of ESS e [MW].
$P^{\mathrm{d\,min}}_e$	Minimum discharge power of a unit of ESS e [MW].
P^{max}_g	Maximum capacity of a unit of generator g [MW].
P^{min}_g	Minimum capacity of a unit of generator g [MW].
r^{dn}_e	Maximum downward ramp rate of a unit of ESS e [MW/h].
r^{dn}_g	Maximum downward ramp rate of a unit of generator g [MW/h].

deterministic static planning for a target year was performed, and long-term uncertainty was not considered.

Extending the deterministic static planning model, [13–15] deal with the multi-year power system planning problem. In [13], a deterministic multi-year generation expansion planning model including UC constraints was developed. Tractability was maintained by using twelve representative days to model the yearly operation. A decomposition approach was proposed in [14] to solve the deterministic multi-year generation planning problem including UC constraints. Results showed that the Dantzig–Wolfe decomposition allows to solve the problem efficiently. In [15], a deterministic multi-year transmission and generation expansion planning under CO_2 and local pollutant emission taxes was proposed. A non-chronological representation of the operation was used. In Refs. [13–15], aspects such as long-term uncertainty and energy storage systems were not included.

Including long-term uncertainty in power system planning represents a challenge in terms of the tractability of the optimization problem [4]. Therefore, decomposition methods and relaxed operational models were used to solve the problem efficiently. A multi-stage stochastic mixed-integer generation expansion planning model

$\bar{r}_g^{\text{prim}}$	Maximum primary reserve of a unit of generator g [MW].
$r_{n,t}^{\text{Rp}}$	Required upward primary reserve at hour t in node n [MW].
$r_{n,t}^{\text{Rs}}$	Required secondary reserve to cover load deviations at hour t in node n [MW].
$\bar{r}_g^{\text{sec}}$	Maximum secondary reserve of a unit of generator g [MW].
r_e^{up}	Maximum upward ramp rate of a unit of ESS e [MW/h].
r_g^{up}	Maximum upward ramp rate of a unit of generator g [MW/h].
$\boldsymbol{U}^0$	$\{0,1\}$ vector indicating if asset is an existing or a candidate asset.
$\boldsymbol{Z}^0$	Vector of installed units of existing assets.
Z_e^{E0}	Installed units of existing ESS e.
$Z_{n,e}^{\text{Emax}}$	Maximum total installed units of ESS e in node n.
Z_g^{G0}	Installed units of existing generator g.
$Z_{n,g}^{\text{Gmax}}$	Maximum total installed units of generator g in node n.
Z_l^{L0}	Installed units of existing line l.
$Z_{n,l}^{\text{Lmax}}$	Maximum total installed lines l in node n.
$\boldsymbol{Z}_n^{\max}$	Vector of maximum total installed units in node n.

Sets

Ω_n	Set of sibling nodes of node n.
B	Set of all system buses.
$\mathcal{E}$	Set of all ESSs.
$\mathcal{E}_0$	Set of all existing ESSs.
$\mathcal{E}_b$	Set of ESSs connected to bus b.
$\mathcal{E}_{\text{new}}$	Set of all candidate ESSs.
G	Set of all generators.
G_0	Set of all existing generators.
G_b	Set of generators connected to bus b.
G_{new}	Set of all candidate generators.
G_{TH}	Set of all thermal generators.
G_{VR}	Set of all variable renewable generators.
K_n	Set of indices of elements in $\mathcal{Z}_n$.
L	Set of all transmission lines.
L_0	Set of all existing transmission lines.
L_{new}	Set of all candidate transmission lines.
L_b^{in}	Set of lines incoming to bus b.
L_b^{out}	Set of lines outgoing from bus b.
$\mathcal{N}$	Set of nodes in scenario tree.
$\mathcal{P}_n$	Set of all predecessors of n in scenario tree.
S	Set of stages in scenario tree.
T	Set of operational time periods.
$\mathcal{X}_n$	Set of feasible operational decisions in node n.
$\mathcal{Z}_n$	Set of feasible total installed units in node n.

considering different environmental policies was proposed in [16]. A simplified representation of the operation without unit commitment constraints was used. The model was solved heuristically using a rolling-horizon approach. In [17], a multi-stage stochastic linear generation and transmission planning model was developed. The model represents the operation by an economic dispatch with ramp rate constraints. Unit commitment constraints such as minimum power, operating times, and reserves were neglected. The Progressive Hedging (PH) algorithm was used to decompose the linear problem through scenarios and reduce solution times. In [18], the multi-stage stochastic transmission and ESS planning problem was dealt with. A method based on Nested Benders Decomposition (NBD) was proposed. The operation was represented by a DC power flow with no consideration of UC constraints. The NBD approach allowed to solve the problem, achieving a solution with a 1.11% optimality gap. However, the methods presented in [16–18] do not cover long-term uncertainty and UC constraints in a unique planning framework. More recently, a stochastic generation and storage planning model with a more detailed operational model was proposed in [19]. To solve the problem using Stochastic Dual Dynamic Integer Programming, uncertainty was assumed to be stage-wise independent. Although the accuracy of the model was improved, the stage-wise independence assumption is rather restrictive [20].

The contributions of this paper are threefold. Firstly, an integrated multi-stage stochastic generation, transmission, and energy storage planning model accounting for short-term flexibility requirements is proposed. The model enables an accurate representation of both short-term operation and long-term uncertainty in a unique planning framework. Thus, the value of investing in flexibility options can be correctly assessed. Due to the significant role that energy storage systems will play in renewable power systems, a general ESS model is proposed to represent the most important characteristics of the different technologies. The resulting planning model is a large-scale mixed-integer programming (MIP). Therefore, the second contribution is the development of a parallel processing solution framework for leveraging high-performance distributed computing in long-term planning under uncertainty. The proposed framework makes use of the novel Column Generation and Sharing decomposition to separate the problem in several smaller subproblems that can be solved in parallel. Thirdly, extensive computational experiments are conducted to study the performance and added value of the proposed method in two aspects. On the one hand, different planning models are compared to analyze the improvement in solution quality when considering both short-term constraints and long-term uncertainty. On the other hand, different state-of-the-art decomposition methods are implemented and compared with the proposed methodology. Results show that the proposed methodology increases the solution speed by 86 % and allows to solve instances that cannot be handled by existing methods.

The framework proposed in this paper differs in several aspects from [19]. Firstly, the model proposed here considers investment also in transmission. Secondly, the operational model for the generators is more accurate, including also operating times constraints originally neglected. Thirdly, the ESS model proposed here is more general including also minimum and maximum charge and discharge power and ramp rates. Moreover, ESS can participate also in reserve provision. Fourthly, the highly restrictive stage-wise independence assumption is not required by the novel solution method.

The remainder of this paper is organized as follows. The multi-stage stochastic generation, transmission, and storage expansion planning model is proposed in Section 2. The details of the parallel solution framework are presented in Section 3. The performance and added value of the solution method are shown through computational experiments in Section 4. In Section 5, the proposed planning framework is used to assess the value of ESSs in the transition to a low-carbon power system for a realistic-size system. Section 6 summarizes the conclusions.

2. Multi-stage stochastic power system planning model

The power system planning problem is a multi-stage stochastic problem as several parameters are uncertain at the moment of decision. Investments can be made in different assets such as generation, transmission, and ESSs. In Section 2.1, a base stochastic planning model is presented, and the treatment of long-term uncertainty is explained. In

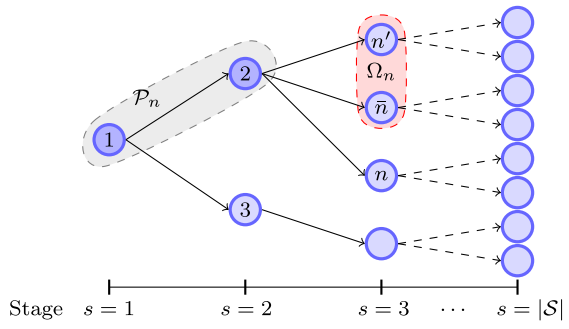

Fig. 1. Scenario tree representation of uncertainties over $|\mathcal{S}|$ decision stages, node n at stage s representing possible state of uncertain parameters.

Section 2.2, the integrated multi-stage stochastic power system planning model with flexibility requirements is developed. The flexibility requirements are modeled through a unit commitment. A general model for ESSs is proposed, with the aim of representing the capabilities and limitations of different ESS technologies. A CO_2 emission constraint is included to enforce the transition to a low-carbon system. This will allow to study the optimal mix of generation, ESSs, and transmission for a cost-effective power system transformation.

2.1. Base multi-stage stochastic power system planning model

The power system planning model is formulated as a multi-stage stochastic problem to account for long-term uncertainty. The scenario-tree-based formulation of [21] serves as a suitable general basis for extension and modification. In the power system planning problem, long-term uncertainty is associated with different parameters, such as fuel prices, investment costs, and future demand. Long-term uncertainty is modeled by a scenario tree over $|\mathcal{S}|$ decision stages,[1] as in Fig. 1. Node n in stage s of the tree constitutes a potential state of the uncertain parameters at stage s. Set $\mathcal{P}_n$ indicates the predecessor nodes of n. Set Ω_n denotes the sibling nodes of n. The root node represents the initial state of the system. In this node, only investment decisions are made. The following nodes consist of an operational phase, followed by an investment phase where the investments for the successor nodes are decided. In this structure, new assets are available to operate one stage after their investment is decided. This multi-stage approach allows to make investment decisions at several points in time and consider the information about the uncertain parameters that is known at the current stage. Short-term uncertainty is not explicitly represented in the multi-stage approach. Nevertheless, the effect of short-term uncertainty on the investment decisions is captured in the model presented in Section 2.2.

Following the scenario-tree-based formulation, the base multi-stage stochastic power system planning model is described by objective function (1) and constraints (2)–(20). Objective function (1) minimizes the expected sum of investment and operational costs. Investments can be made in generating units, transmission lines, and energy storage systems. Constraints (2)–(4) are the non-anticipativity constraints [22]. They relate the investments made in the different assets at the predecessor nodes of n with the total installed units in node n. A delay of one stage between the investment decision and the availability of the asset is considered. Constraints (2)–(4) also limit the number of total installed units in node n. The transmission system is represented by a transportation model [23]. Therefore, constraints pertaining to the transmission network are (5)–(6), in addition to the non-anticipativity constraint on transmission investments (4). Constraint (5) ensures the power balance at each bus of the system for each hour of the operational stage. Constraint (6) limits maximum flows through the transmission lines to the respective installed capacity. The simplified model of the generators considers only the maximum generation limit for thermal generators, given by (7). The power of RES generators is limited by the installed units and the availability of the primary source, as shown in (8). The simplified model for ESSs considers the energy balance (9), minimum and maximum energy storage levels (10), and maximum charge and discharge power constraints (11). Constraint (10) allows to limit the energy level of the energy storage to prevent accelerated degradation. For some battery energy storage technologies, it is suggested to avoid operating at a very low or very high state of charge to prevent premature aging [24,25].

$$\min \sum_{n\in\mathcal{N}} \phi_n \Big[\sum_{g\in G_{\text{new}}} c^{\text{inv}}_{n,g} I^{\text{G}}_{n,g} + \sum_{e\in\mathcal{E}_{\text{new}}} c^{\text{inv}}_{n,e} I^{\text{E}}_{n,e} + \sum_{l\in L_{\text{new}}} c^{\text{inv}}_{n,l} I^{\text{L}}_{n,l} + \sum_{t\in T} \Big(\sum_{g\in G} c^{\text{var}}_{n,g} P_{n,t,g} + \sum_{e\in\mathcal{E}} c^{\text{var}}_{n,e} P^{\text{d}}_{n,t,e} + \sum_{b\in B} c^{\text{UD}}_{n} L^{\text{S}}_{n,t,b} \Big) \Big] \tag{1}$$

$$\text{s.t.:} \quad Z^{\text{G}}_{n,g} \le \sum_{h\in\mathcal{P}_n} I^{\text{G}}_{h,g} \le Z^{\text{Gmax}}_{n,g} \qquad \forall n, \ \forall g \in G \tag{2}$$

$$Z^{\text{E}}_{n,e} \le \sum_{h\in\mathcal{P}_n} I^{\text{E}}_{h,e} \le Z^{\text{Emax}}_{n,e} \qquad \forall n, \ \forall e \in \mathcal{E} \tag{3}$$

$$Z^{\text{L}}_{n,l} \le \sum_{h\in\mathcal{P}_n} I^{\text{L}}_{h,l} \le Z^{\text{Lmax}}_{n,l} \qquad \forall n, \ \forall l \in L \tag{4}$$

$$\sum_{g\in G_b} P_{n,t,g} + \sum_{e\in\mathcal{E}_b} (P^{\text{d}}_{n,t,e} - P^{\text{c}}_{n,t,e}) + \sum_{l\in L^{\text{in}}_b} F_{n,t,l} - \sum_{l\in L^{\text{out}}_b} F_{n,t,l} + L^{\text{S}}_{n,t,b} = d_{n,t,b} \qquad \forall n,t,b \tag{5}$$

$$-F^{\max}_l Z^{\text{L}}_{n,l} \le F_{n,t,l} \le F^{\max}_l Z^{\text{L}}_{n,l} \qquad \forall n,t,l \tag{6}$$

$$P_{n,t,g} \le Z^{G}_{n,g} P^{\max}_g \qquad \forall n,t, \ \forall g \in G_{\text{TH}} \tag{7}$$

$$P_{n,t,g} \le Z^{\text{G}}_{n,g} P^{\max}_g \alpha^{\text{avail}}_{n,t,g} \qquad \forall n,t, \ \forall g \in G_{\text{VR}} \tag{8}$$

$$E_{n,t,e} = (1-\eta^{\text{sd}}_e) E_{n,t-1,e} + \eta^{\text{c}}_e P^{\text{c}}_{n,t,e} - P^{\text{d}}_{n,t,e}/\eta^{\text{d}}_e \qquad \forall n,t,e \tag{9}$$

$$E^{\min}_e Z^{\text{E}}_{n,e} \le E_{n,t,e} \le E^{\max}_e Z^{\text{E}}_{n,e} \qquad \forall n,t,e \tag{10}$$

$$P^{\text{d}}_{n,t,e} \le P^{\text{d}^{\max}}_e Z^{\text{E}}_{n,e}, P^{\text{c}}_{n,t,e} \le P^{\text{c}^{\max}}_e Z^{\text{E}}_{n,e} \qquad \forall n,t,e \tag{11}$$

$$Z^{\text{G}}_{n,g} = Z^{\text{G0}}_g \qquad \forall n, \ \forall g \in G_0 \tag{12}$$

$$Z^{\text{E}}_{n,e} = Z^{\text{E0}}_e \qquad \forall n, \ \forall e \in \mathcal{E}_0 \tag{13}$$

$$Z^{\text{L}}_{n,l} = Z^{\text{L0}}_l \qquad \forall n, \ \forall l \in L_0 \tag{14}$$

$$I^{\text{G}}_{1,g} = Z^{\text{G0}}_g \qquad \forall g \in G_0 \tag{15}$$

$$I^{\text{E}}_{1,e} = Z^{\text{E0}}_e \qquad \forall e \in \mathcal{E}_0 \tag{16}$$

$$I^{\text{L}}_{1,l} = Z^{\text{L0}}_l \qquad \forall l \in L_0 \tag{17}$$

$$I^{\text{G}}_{n,g} = 0 \qquad \forall n > 1, \ \forall g \in G_0 \tag{18}$$

$$I^{\text{E}}_{n,e} = 0 \qquad \forall n > 1, \ \forall e \in \mathcal{E}_0 \tag{19}$$

$$I^{\text{L}}_{n,l} = 0 \qquad \forall n > 1, \ \forall l \in L_0 \tag{20}$$

Constraints (12)–(20) are introduced to model existing assets. Constraints (12)–(14) set the number of installed units of existing assets in every node n to the number of existing units at the root node. To maintain the consistency of the model, constraints (15)–(17) set the number of additional units installed at the root node to the number of existing units. After the root node, no additional units can be installed for existing assets, as stated by (18)–(20).

2.2. Integrated multi-stage stochastic power system planning with flexibility requirements

The proposed integrated planning model improves the model presented in Section 2.1 by including a detailed representation of the

[1] A stage could represent a year or periods comprising multiple years.

hourly operation in the form of unit commitment constraints. This is the first of the scientific contributions listed in Section 1. The hourly representation of the operation allows capturing demand and renewable generation variability. The unit commitment constraints capture the ability of generation and energy storage technologies to respond to this variability. Also, the operating reserve constraints of the unit commitment model allow to represent short-term uncertainty. The flexibility requirements of power systems with high penetration of renewable energy can so be quantified accurately. Thus, the value of investing in flexibility options can be assessed. The flexibility options considered in the planning model are energy storage systems and open-cycle gas turbines. The UC model developed here comprises system, generator, and ESS constraints. The system constraints cover the power balance, operating reserve requirements, and CO_2 emission limits. The operational model of generators includes minimum and maximum output power, minimum operation time intervals, ramping rates, and reserve capability constraints. In the case of the ESS, the operation model not only considers the energy balance constraint but also minimum and maximum charge and discharge power, ramping rates, and operating reserve constraints, improving the ESS representation used for planning purposes so far [18,19].

In the proposed integrated model, long-term uncertainty is represented by the scenario-tree-based formulation of Section 2.1. The impact of short-term uncertainty in investment decisions is implicitly modeled by operating reserve requirements. Operating reserves are allocated to manage uncertainty in the short-term operation [26,27]. These reserve requirements will encourage investing in flexibility options to cope with imbalances caused by short-term load and RES uncertainty. Moreover, the detailed representation of the hourly operation eliminates the need for an explicit model of short-term uncertainties. In [28], a power system planning model considering short-term uncertainty and a counterpart without short-term uncertainty were compared. Results show that the difference in total costs and investment decisions obtained by both models is almost negligible if the short-term operational constraints are included and a sufficient number of representative days is used. This is because the flexibility options installed to cope with hourly variability of demand and renewable generation will provide most of the flexibility required to respond to imbalances caused by short-term uncertainty.

Incorporating the UC in long-term planning leads to a high-dimensional problem. To reduce the size of the problem, a clustered UC model is used. A similar formulation was used in [5] for deterministic static planning. As opposed to [5], the proposed model includes long-term uncertainty, ESSs, and the transmission network. In the clustered UC, generating units with similar characteristics are combined into a single equivalent generator with multiple units. Therefore, investment and commitment decisions become integer variables. In what follows, the system, generator, and ESS constraints are introduced.

2.2.1. System constraints

In addition to the power balance (5), constraints such as operating reserve and emission limits are included. Constraints (21)–(23) represent the primary and secondary reserve requirements. These reserves can be deployed in real time to manage imbalances due to short-term uncertainty of generation and load [26,27]. Primary reserve is required to limit frequency excursions after a disturbance or event [27,29]. In this work, primary reserve is determined such that a sudden loss of generated power can be covered, thus upward primary reserve is being quantified. Primary reserve resources must be fully operational within seconds. Secondary reserve aims to restore the system frequency and to release the primary reserve. Secondary reserve must be fully available within minutes [27]. Including operating reserve requirements allows to model short-term uncertainty implicitly. Constraint (24) imposes a maximum limit on the CO_2 emissions from thermal generation and ESSs. The CO_2 emission coefficient γ_e is greater than zero for energy storage technologies that use fossil fuels, such as compressed air energy storage.

$$\sum_{g\in G_{\mathrm{TH}}} R^{\mathrm{prim}}_{n,t,g} + \sum_{e\in\mathcal{E}} R^{\mathrm{prim}}_{n,t,e} \geq r^{\mathrm{Rp}}_{n,t} \qquad \forall n,t \tag{21}$$

$$\sum_{g\in G_{\mathrm{TH}}} \overline{R}^{\mathrm{sec}}_{n,t,g} + \sum_{e\in\mathcal{E}} \overline{R}^{\mathrm{sec}}_{n,t,e} - \sum_{g\in G_{\mathrm{VR}}} \alpha^{\mathrm{R}} P_{n,t,g} \geq r^{\mathrm{Rs}}_{n,t} \qquad \forall n,t \tag{22}$$

$$\sum_{g\in G_{\mathrm{TH}}} \underline{R}^{\mathrm{sec}}_{n,t,g} + \sum_{e\in\mathcal{E}} \underline{R}^{\mathrm{sec}}_{n,t,e} - \sum_{g\in G_{\mathrm{VR}}} \alpha^{\mathrm{R}} P_{n,t,g} \geq r^{\mathrm{Rs}}_{n,t} \qquad \forall n,t \tag{23}$$

$$\sum_{t\in T}\Big(\sum_{g\in G_{\mathrm{TH}}} \gamma_g P_{n,t,g} + \sum_{e\in\mathcal{E}} \gamma_e P^{\mathrm{d}}_{n,t,e}\Big) \leq \Gamma^{\max}_{n} \qquad \forall n \tag{24}$$

2.2.2. Generating unit constraints

The operation of the generators is modeled by (25)–(35). Maximum and minimum generation limits are imposed by constraints (25) and (26), respectively. Constraints (27)–(29) represent the maximum primary and secondary reserve that can be provided by generators. These limits consider also that reserve capacity is limited by the ramping rates. Start-up and shut-down variables are included to represent start-up costs and to constrain minimum operation times. These variables are related through the unit commitment state equation (30). The ramping limits of generators are imposed by (31) and (32). It is assumed that, when a unit is started, its output is limited by $P^{\min}_{g}$. When a unit is shut down, it can immediately reduce its power from $P^{\min}_{g}$ to zero. Constraints (33) and (34) impose minimum operating times. Constraint (35) restricts the number of units that can be committed to the number of units that have been built.

$$P_{n,t,g} + R^{\mathrm{prim}}_{n,t,g} + \overline{R}^{\mathrm{sec}}_{n,t,g} \leq U_{n,t,g} P^{\max}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{25}$$

$$U_{n,t,g} P^{\min}_{g} + \underline{R}^{\mathrm{sec}}_{n,t,g} \leq P_{n,t,g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{26}$$

$$R^{\mathrm{prim}}_{n,t,g} \leq U_{n,t,g} \overline{r}^{\mathrm{prim}}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{27}$$

$$\overline{R}^{\mathrm{sec}}_{n,t,g} \leq U_{n,t,g} \overline{r}^{\mathrm{sec}}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{28}$$

$$\underline{R}^{\mathrm{sec}}_{n,t,g} \leq U_{n,t,g} \overline{r}^{\mathrm{sec}}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{29}$$

$$U_{n,t,g} = U_{n,t-1,g} + S_{n,t,g} - D_{n,t,g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{30}$$

$$P_{n,t,g} - P_{n,t-1,g} \leq U_{n,t-1,g} r^{\mathrm{up}}_{g} + S_{n,t,g} P^{\min}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{31}$$

$$P_{n,t-1,g} - P_{n,t,g} \leq U_{n,t-1,g} r^{\mathrm{dn}}_{g} + D_{n,t,g} P^{\min}_{g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{32}$$

$$U_{n,t,g} \geq \sum_{t'=t-\tau^{\mathrm{on}}_{g}}^{t} S_{n,t',g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{33}$$

$$Z^{\mathrm{G}}_{n,g} - U_{n,t,g} \geq \sum_{t'=t-\tau^{\mathrm{off}}_{g}}^{t} D_{n,t',g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{34}$$

$$U_{n,t,g} \leq Z^{\mathrm{G}}_{n,g} \qquad \forall n,t,\quad \forall g \in G_{\mathrm{TH}} \tag{35}$$

2.2.3. Energy storage system constraints

The operation of ESSs is modeled by (36)–(47), in addition to the energy balance (9) and storage capacity constraint (10). Maximum and minimum discharge and charge power are represented by (36) and (37), respectively. Minimum power is required for some technologies such as pumped hydro storage or compressed air energy storage. Ramping limits are imposed by (38) and (39). Constraints (40) and (41) ensure that there is sufficient energy in the storage system to provide operating reserve during the required time. It is stated that the energy at the end of the reserve provision time must be within the capacity limits of the storage system [30]. Not including this can result in an overestimation of the capabilities of ESS. It is assumed that the reserve provision time does not exceed the one-hour time step of the model. The power that an ESS can provide for reserve depends on the charging power and discharging power of the ESS, as shown in (42) and (43). The reserve provision of the ESS is limited by its maximum ramp rates, as shown in (44)–(46). Constraint (47) limits the number of storage units that can operate to the number of units that have been built.

$$P^{\mathrm{d}\min}_{e} U^{\mathrm{Ed}}_{n,t,e} \leq P^{\mathrm{d}}_{n,t,e} \leq P^{\mathrm{d}\max}_{e} U^{\mathrm{Ed}}_{n,t,e} \qquad \forall n,t,e \tag{36}$$

$$P_e^{\mathrm{c}\min} U_{n,t,e}^{\mathrm{Ec}} \le P_{n,t,e}^{\mathrm{c}} \le P_e^{\mathrm{c}\max} U_{n,t,e}^{\mathrm{Ec}} \quad \forall n,t,e \quad (37)$$

$$P_{n,t,e}^{\mathrm{d}} - P_{n,t,e}^{\mathrm{c}} - (P_{n,t-1,e}^{\mathrm{d}} - P_{n,t-1,e}^{\mathrm{c}}) \le r_e^{\mathrm{up}} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (38)$$

$$P_{n,t-1,e}^{\mathrm{d}} - P_{n,t-1,e}^{\mathrm{c}} - (P_{n,t,e}^{\mathrm{d}} - P_{n,t,e}^{\mathrm{c}}) \le r_e^{\mathrm{dn}} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (39)$$

$$(1-\eta_e^{\mathrm{sd}}\beta^{\mathrm{Rs}})E_{n,t-1,e} + (\eta_e^{\mathrm{c}} P_{n,t,e}^{\mathrm{c}} - P_{n,t,e}^{\mathrm{d}}/\eta_e^{\mathrm{d}})\beta^{\mathrm{Rs}} - 1/\eta_e^{\mathrm{d}}(\beta^{\mathrm{Rp}} R_{n,t,e}^{\mathrm{prim}} + \beta^{\mathrm{Rs}} \overline{R}_{n,t,e}^{\mathrm{sec}}) \ge E_e^{\min} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (40)$$

$$(1-\eta_e^{\mathrm{sd}}\beta^{\mathrm{Rs}})E_{n,t-1,e} + (\eta_e^{\mathrm{c}} P_{n,t,e}^{\mathrm{c}} - P_{n,t,e}^{\mathrm{d}}/\eta_e^{\mathrm{d}})\beta^{\mathrm{Rs}} + \eta_e^{\mathrm{c}}\beta^{\mathrm{Rs}} \underline{R}_{n,t,e}^{\mathrm{sec}} \le E_e^{\max} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (41)$$

$$\overline{R}_{n,t,e}^{\mathrm{sec}} + R_{n,t,e}^{\mathrm{prim}} \le P_e^{\mathrm{d}\max} Z_{n,e}^{\mathrm{E}} - P_{n,t,e}^{\mathrm{d}} + P_{n,t,e}^{\mathrm{c}} \quad \forall n,t,e \quad (42)$$

$$\underline{R}_{n,t,e}^{\mathrm{sec}} \le P_e^{\mathrm{c}\max} Z_{n,e}^{\mathrm{E}} - P_{n,t,e}^{\mathrm{c}} + P_{n,t,e}^{\mathrm{d}} \quad \forall n,t,e \quad (43)$$

$$R_{n,t,e}^{\mathrm{prim}} \le r_e^{\mathrm{up}} \Delta^{\mathrm{Rp}} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (44)$$

$$\overline{R}_{n,t,e}^{\mathrm{sec}} \le r_e^{\mathrm{up}} \Delta^{\mathrm{Rs}} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (45)$$

$$\underline{R}_{n,t,e}^{\mathrm{sec}} \le r_e^{\mathrm{dn}} \Delta^{\mathrm{Rs}} Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (46)$$

$$U_{n,t,e}^{\mathrm{Ec}} + U_{n,t,e}^{\mathrm{Ed}} \le Z_{n,e}^{\mathrm{E}} \quad \forall n,t,e \quad (47)$$

2.2.4. Integrated multi-stage stochastic power system planning model

The multi-stage stochastic power system planning model with flexibility constraints is described by problem (48). The objective function (48a) now includes the start-up costs. Variables $I_{n,g}^{\mathrm{G}}$, $I_{n,e}^{\mathrm{E}}$, $I_{n,l}^{\mathrm{L}}$, $Z_{n,g}^{\mathrm{G}}$, $Z_{n,e}^{\mathrm{E}}$, $Z_{n,l}^{\mathrm{L}}$, $U_{n,t,g}$, $U_{n,t,e}^{\mathrm{Ec}}$, $U_{n,t,e}^{\mathrm{Ed}}$, $S_{n,t,g}$, and $D_{n,t,g}$ are integers, variable $F_{n,t,l}$ is a continuous variable, the remaining ones are non-negative continuous variables.

$$\min \sum_{n\in\mathcal{N}} \phi_n \Big[\sum_{g\in G_{\mathrm{new}}} c_{n,g}^{\mathrm{inv}} I_{n,g}^{\mathrm{G}} + \sum_{e\in\mathcal{E}_{\mathrm{new}}} c_{n,e}^{\mathrm{inv}} I_{n,e}^{\mathrm{E}} + \sum_{l\in L_{\mathrm{new}}} c_{n,l}^{\mathrm{inv}} I_{n,l}^{\mathrm{L}} + \sum_{t\in T}\Big(\sum_{g\in G} \Big(c_{n,g}^{\mathrm{var}} P_{n,t,g} + c_g^{\mathrm{s}} S_{n,t,g} \Big) + \sum_{e\in\mathcal{E}} c_{n,e}^{\mathrm{var}} P_{n,t,e}^{\mathrm{d}} + \sum_{b\in B} c_n^{\mathrm{UD}} L_{n,t,b}^{\mathrm{S}} \Big) \Big] \quad (48a)$$

s.t.: (2)–(6), (8)–(10), (12)–(20), (21)–(47). (48b)

3. Parallel solution framework

The multi-stage stochastic planning problem presented in Section 2.2 is a large-scale mixed-integer programming problem. Even corresponding deterministic models are known to be challenging to solve [14]. Furthermore, the multi-stage stochastic problem becomes easily intractable. Even for small-size systems, the model could be too large to be solved by existing commercial solvers. Therefore, a well-suited solution method is required.

The proposed parallel solution framework is depicted in Fig. 2. First, problem (48) is decomposed into a master problem and $|\mathcal{N}|$ subproblems using the Dantzig–Wolfe reformulation described in Appendix A. Following the reformulation, a parallel solution using distributed computing is enabled by the novel Column Generation & Sharing (CG&S) method. There, the master problem and subproblems are constructed on the master and worker processors, respectively. An optimal solution to problem (48) is found iteratively by the workers coordinated by the master processor. The main innovative feature is the Column Sharing (CS) procedure, which leads to a significant speed-up by sharing information among the subproblems. The proposed CG&S method extends the standard Column Generation (CG) [31] algorithm. It allows to exploit not only the diagonal block structure of problem (48), but also the scenario tree structure. The diagonal block structure is exploited by the decomposition, while the scenario tree structure is exploited by the novel Column Sharing procedure.

The proposed parallel solution framework is the second scientific contribution listed in Section 1. In Section 3.1, the CG&S method is presented, and the distributed computing implementation is discussed. In Section 3.2, the novel Column Sharing procedure is proposed. Section 3.3 discusses the convergence guarantee of the CG&S method.

3.1. Column generation and sharing method

Fig. 2 depicts the Column Generation and Sharing method within the parallel solution framework. Before starting the CG&S method, problem (48) is reformulated using the Dantzig–Wolfe decomposition. The details of the proposed reformulation are described in Appendix A. By this mean, a master problem (MP) and $|\mathcal{N}|$ subproblems SP_n are obtained. Then, the master problem and the subproblems are constructed in the master and worker processors, respectively. Over the course of the CG&S method, problem (48) is solved in a distributed fashion by the workers through the coordination of the master processor. Information flows between the processors as depicted by the dashed lines in Fig. 2. The openMPI implementation of the Message Passing Interface is used to communicate between the processors.

The CG&S method starts by solving the linear programming (LP) relaxation of the master problem in the master processor. In this step, dual prices $(\boldsymbol{\pi}_n, \mu_n)$ are obtained for each node n of the scenario tree. The dual prices are sent to the worker processors, as shown in Fig. 2. In the workers, the subproblems are solved to generate columns with minimum reduced cost z_n^{sp}, defined in Appendix A. A column is defined as a pair of feasible total installed units and its associated optimal operation for node n, denoted by $\{\widehat{\boldsymbol{Z}}_n, \widehat{\boldsymbol{X}}_n\}$. The columns generated by the subproblems and their reduced cost z_n^{sp} are sent back to the master processor.

In a next step, the novel Column Sharing procedure is applied. Here, columns are shared among the workers to generate additional columns $\{\widehat{\boldsymbol{Z}}_n^{\mathrm{s}}, \widehat{\boldsymbol{X}}_n^{\mathrm{s}}\}$ that improve the efficiency of the algorithm. The columns generated by the Column Sharing procedure are sent back to the master processor, as shown in Fig. 2. Afterwards, the columns are added to the MP. In the case of the columns generated by the subproblems, only the ones with negative reduced cost z_n^{sp} are added to the MP. In the case of the columns generated by the CS procedure, all of them are added to the MP. Then, the LP gap is computed as in [32], and the termination criterion is checked. If columns with negative reduced cost exist, a new iteration is performed. If not, the termination criterion is met[2] with $gap_{\mathrm{LP}} = 0$ [32].

Since the LP relaxation of the master problem is being solved, fractional values can be obtained. Therefore, integrality of the solution is enforced after the termination criterion is met, by solving the Integer MP. Then, the MIP gap is computed, and the algorithm terminates.[3]

3.2. Novel column sharing procedure

The aim of the Column Sharing procedure is to overcome known convergence issues of the CG approach [31]. In general, convergence of the CG approach could be improved if more information were added to the master problem [32]. In this sense, the idea is to exploit the scenario tree structure to find feasible and relevant columns with less computational effort. The novel CS procedure generates new columns by sharing information among the subproblems, as shown in Fig. 2. By adding new columns, more information is added to the MP, and the price information for the subproblems is improved.

The Column Sharing procedure allows to eliminate convergence issues of the CG approach related to the 'heading-in"[4] and "tailing-off"[5] effects. A CS procedure was proposed in [33] for recombining scenario trees. The CS procedure proposed here is improved in that it applies to recombining and non-recombining scenario trees.

[2] The algorithm could be terminated when the gap_{LP} is below a threshold.

[3] If gap_{MIP} is not in the required range, branch-and-price should be used.

[4] The 'heading-in" effect occurs in the first iterations when there is a small number of columns in the master problem. Thus, the obtained prices are of poor quality. This results in irrelevant columns and a poor lower bound [31].

[5] The "tailing-off" effect occurs when the algorithm is close to the optimum and little progress per iteration is obtained [31].

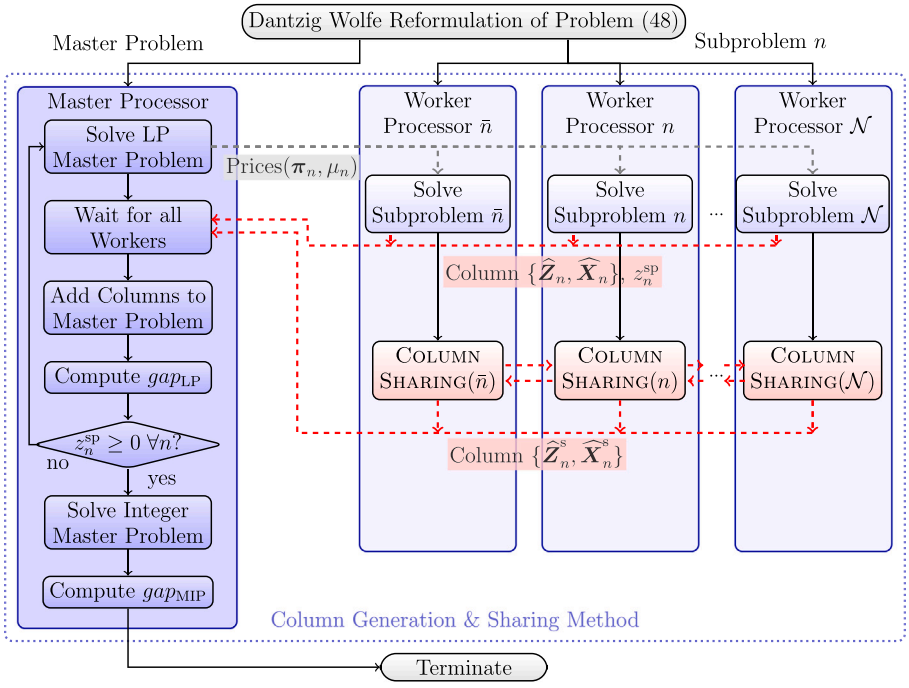

Fig. 2. Diagram of parallel solution framework.

The novel CS procedure, described by Algorithm 1, works in the following way. Once the subproblem for node n has been solved, a column $\{\widehat{Z}_n, \widehat{X}_n\}$ is obtained, as described in Section 3.1. The vector $\widehat{Z}_n$ contains the total installed units for node n, and it is also a feasible solution for the installed units of the subproblems for its sibling nodes, denoted by the set Ω_n. Then, the vector $\widehat{Z}_n$ is shared with each sibling node $\bar{n} \in \Omega_n$. Reciprocally, node n receives vector $\widehat{Z}_{\bar{n}}$ from each sibling node $\bar{n}$. For node n, an additional column can be obtained by solving the operation with installed units fixed to $\widehat{Z}_{\bar{n}}$. In this way, $|\Omega_n|$ new columns are obtained per node n. These additional columns can be generated with reduced effort since solving the operation with fixed capacities is significantly faster than solving the subproblem SP_n, as defined in Appendix A. Moreover, these new columns are relevant to the master problem because they intrinsically meet the non-anticipativity constraints (2)–(4).

Algorithm 1 Column sharing procedure.

1: **procedure** ColumnSharing(n)
2: **for** $\bar{n} \in \Omega_n$ **do**
3: Send $\widehat{Z}_n$ to node $\bar{n}$, receive $\widehat{Z}_{\bar{n}}$ from node $\bar{n}$
4: Compute operation $\widehat{X}_n^{\text{s}}$ for installed units $\widehat{Z}_n^{\text{s}} = \widehat{Z}_{\bar{n}}$
5: Send column $\{\widehat{Z}_n^{\text{s}}, \widehat{X}_n^{\text{s}}\}$ to Master Processor

3.3. Convergence of column generation and sharing method

The convergence proof of the proposed improved Column Generation and Sharing algorithm is sketched in the following. In the case where the decision variables of the master problem are continuous, convergence of the algorithm can be proven following the convergence proof for the standard CG. The number of columns that can be generated is finite because set $\mathcal{Z}_n$, defined in Appendix A, is a bounded integer polyhedron [31]. In each iteration, a column is obtained by solving subproblem SP_n, and $|\Omega_n|$ columns are obtained by the CS procedure, for each node n. At the end of the iteration, the algorithm either terminates, or there is at least one column $\{\widehat{Z}_n, \widehat{X}_n\}$ with negative reduced cost z_n^{sp}. This column must be different from the ones that had been added to the MP in previous iterations because no column in the MP can have negative reduced cost [32]. Therefore, after a finite number of iterations, no more columns with negative reduced cost z_n^{sp} exist, and the algorithm terminates. Convergence in the case where the decision variables of the MP are binary can be proven by branch-and-price with proper branching rules [31].

4. Validation of performance

In this section, the performance and added value of the proposed planning framework is validated in two aspects. Firstly, the improvement in solution quality when considering both UC constraints and long-term uncertainty is analyzed in Section 4.2. The integrated planning model of Section 2.2 is compared with different state-of-the-art planning models. Secondly, the computational performance of the proposed Column Generation and Sharing method is studied in Section 4.3. The Column Generation and Sharing method is compared with different state-of-the-art solution methods for solving the integrated planning model of Section 2.2. The validation of performance and benchmarking of the proposed method against existing state-of-the-art methods is the third contribution listed in Section 1.

The computational experiments of Sections 4.2 and 4.3 are based on the NREL 118-bus system [34] described in Section 4.1. The different models and solution methods were implemented in Java and solved using CPLEX 12.8 on the High Performance Computing cluster of TU Berlin. The cluster is composed of 132 nodes, equipped with two 2.67 GHz quad-core Intel Xeon E5-2630 v4 processors and 256 GB RAM.

4.1. Test system

The computational experiments are based on the NREL 118-bus system [34] for the period 2016 to 2040. The NREL 118-bus system, shown in Fig. 3, can be considered as a medium-size system [34]. It was selected due to its higher data resolution and higher RES penetration compared to the IEEE 118-bus system.

The NREL 118-bus system consists of 118 buses, 186 lines, and 327 generation units. Hydro and geothermal generators are not considered.

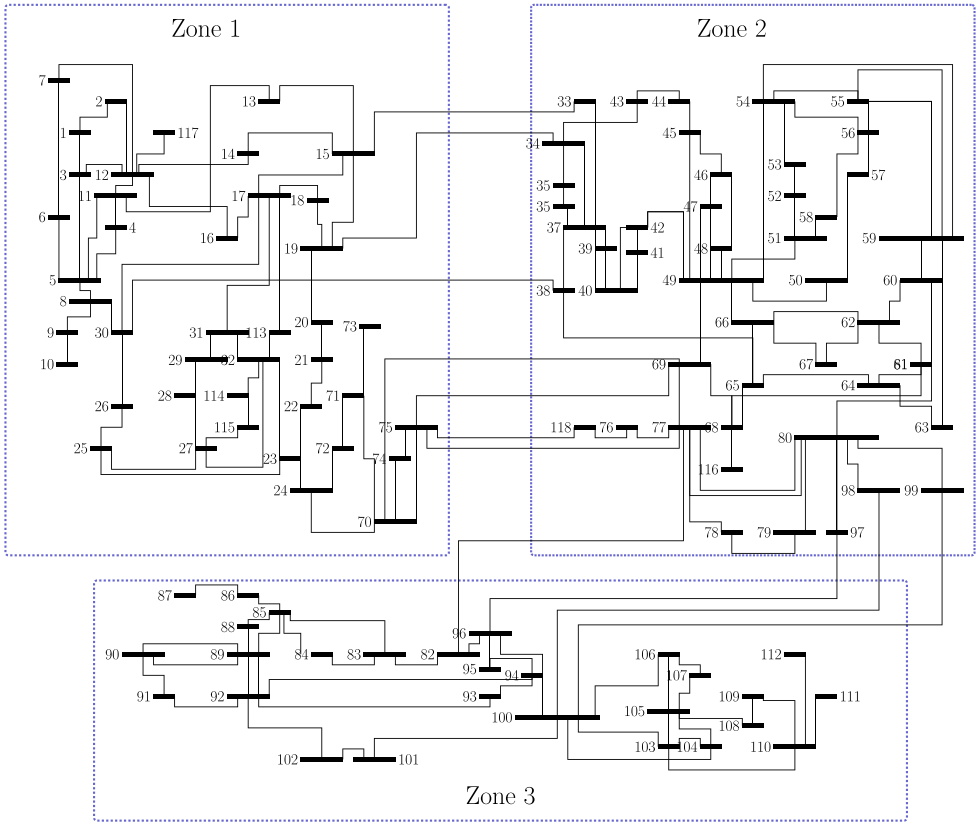

Fig. 3. One-line diagram of NREL 118-bus system.

Steam turbines are supplied by coal. Existing generators are clustered by technology and location, resulting in 46 generation plants. The investment options are open-cycle gas turbines, combined cycle gas turbines, solar generation, pumped hydro storage (PHS), compressed air energy storage (CAES), batteries, and transmission lines. Among these investment candidates, open-cycle gas turbines and ESSs are considered as flexibility options. Future investment in wind is considered given. In total, 18 generation plants, 10 ESSs, and 10 lines are candidates for investment. Load and solar profiles from [34] are used. Generator parameters and investment costs are obtained from [35]. Technical parameters of energy storage technologies and investment costs are obtained from [36,37]. Based on [35,36], investment costs of renewable generation, batteries, and CAES show a decreasing trend. Fuel prices are obtained from [38]. In each stage, a CO_2 emission limit is imposed with the aim of reducing emissions by 80 % in 2040 with respect to the initial year. Further details of the input data are provided in Appendix D.

The scenario tree is an input to the methodology. To illustrate the performance of the proposed method, uncertainty is considered in gas price and demand growth. The scenarios are generated following [39,40]. Demand and gas price are modeled as two correlated geometric Brownian motion processes. The scenario tree is generated using the moment matching method with three branches per node, as described in Appendix B. No further scenario reduction is applied. Up to five stages comprising five years each are considered. The first stage represents the root node and corresponds to the period 2016 to 2020. Each stage is represented by typical days or weeks.

4.2. Impact of unit commitment constraints and uncertainty

For the purpose of analyzing the impact of UC constraints and long-term uncertainty, three different planning models were considered:

- SP-noUC: Stochastic planning without UC constraints [18].
- DP-UC: Deterministic planning with UC constraints [14]. Expected values are used for fuel prices and demand growth.
- SP-UC: Proposed stochastic planning model with UC constraints.

Since the aim is to study the accuracy of the modeling rather than the computational performance, the NREL-118 system was reduced to a three-region representation.[6] The planning horizon comprises five stages. Each stage was represented by four typical weeks to account for the seasonality of load and RES. For each planning model, the optimal investment plan was obtained, and the investment cost was computed. Afterwards, the expected operation cost of each investment plan was computed by simulating the operation with UC constraints for each node of the scenario tree.

The results are given in Table 1 and Fig. 4. Table 1 shows the expected investment, operation and total costs, unserved load, and RES curtailment. Fig. 4 shows the expected installed capacities for each planning model. Stage 1 is omitted as it represents the initial state of the system.

The stochastic planning model without UC constraints overestimated the capability of the system to integrate RES and underestimated the value of flexibility options. The investment in solar energy was increased, and the investment in flexibility options, specifically in batteries, was reduced, as shown in Fig. 4. The lack of flexibility led to unserved energy, higher RES curtailment, and 47 % higher total system cost compared to the proposed stochastic planning model with UC constraints, as shown in Table 1.

The solution obtained by the deterministic planning model with UC constraints is not well adapted to the different possible scenarios. From stage 3 on, the deterministic model invested less capacity than

[6] The three regions are defined as in [34], and shown in Fig. 3. Generators and demand are aggregated by region, and intra-regional transmission constraints are neglected.

Table 1
Expected costs and operational results in validation case of Section 4.2.

Model	Investment cost [billion Euro]	Operation cost [billion Euro]	Total cost [billion Euro]	Unserved load [TWh]	RES Curtailment [TWh]
Stochastic planning without UC constraints (SP-noUC)	21	61.8	82.8	5.4	236.8
Deterministic planning with UC constraints (DP-UC)	18.3	43.4	61.8	1.4	143.2
Stochastic planning with UC constraints (SP-UC)	18.8	37.4	56.2	0.01	166.9

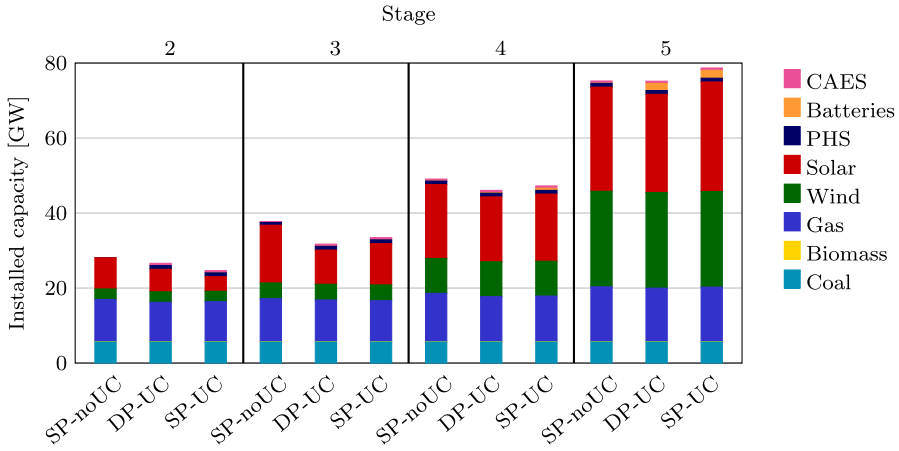

Fig. 4. Expected installed capacities in validation case of Section 4.2.

the stochastic planning model with UC constraints, as shown in Fig. 4. Thus, it could not cope with scenarios of high demand. This resulted in unserved load. The system cost was 9 % higher than for the stochastic planning model with UC constraints, as also shown in Table 1.

In summary, the results show that neglecting UC constraints or long-term uncertainty leads to suboptimal planning decisions. The proposed integrated model presented in Section 2.2 significantly improves the solution quality compared with the state-of-the-art models by including UC constraints and long-term uncertainty in a unique planning framework.

4.3. Computational performance

In this section, the computational performance of the novel CG&S algorithm for solving the multi-stage stochastic planning model with UC constraints proposed in Section 2.2 is studied. To do so, the proposed Column Generation and Sharing method is compared with different state-of-the-art solution methods such as the PH algorithm, Nested Benders decomposition, and the Column Generation algorithm. Two variants of the CG algorithm are considered. These variants implement state-of-the-art stabilization techniques with the aim of improving the CG algorithm. A discussion on stabilization techniques for the CG algorithm is provided in Appendix C. In summary, the following solution methods were implemented using distributed computing and used to solve problem (48) of Section 2.2:

- Undecomposed model (UD): Problem (48) is solved with CPLEX 12.8 using 20 CPUs.
- CG with interior point stabilization (CG-ip): Problem (48) is decomposed by nodes of the scenario tree and solved with the CG approach. To improve the convergence, the interior point stabilization [14,41] is used.
- CG with dual smoothing (CG-ds): Problem (48) is decomposed by nodes of the scenario tree and solved with the CG approach. To improve the convergence, the auto-adaptive Wentges dual price smoothing [42] is used.
- Column Generation and Sharing (CG&S): Problem (48) is decomposed by nodes of the scenario tree and solved using the proposed Column Generation and Sharing algorithm. The interior point method is used to solve the master problem of the CG&S method.
- Progressive Hedging[7] (PH): Problem (48) is decomposed by scenarios and solved with PH [17]. The variable fixing heuristic is used to avoid cyclic behavior [48].
- Nested Benders (NBD): Problem (48) is decomposed by nodes and solved using the NBD [18]. To improve the convergence, the strengthened Benders' cuts [19] are used.

Four instances were constructed by increasing the number of stages $|S|$ and nodes $|\mathcal{N}|$ in the scenario tree. No reduction was applied to the NREL 118-Bus system. Each stage was represented by one typical day to avoid memory issues for the UD model. A MIP gap of 0.5 % was required.

Table 2 shows the numbers of variables and constraints of the subproblems of each method. The total size of the problem is shown for the UD model. The size of the UD model grows exponentially with the number of stages $|S|$. The size of the PH subproblems grows linearly with the number of stages. This growth can become an issue for larger problems. The size of the subproblems of NBD, CG-ip, CG-ds, and CG&S remains constant since these methods decompose by nodes of the scenario tree. The number of constraints in the NBD subproblems will increase in each iteration as new cuts are added [18].

Table 3 summarizes the computational performance of the solution methods. The solution time and number of iterations are shown. For the methods where no solution within the 0.5 % gap was found within 24 h, the resulting optimality gap is given. The results clearly show the benefit of using the novel Column Generation and Sharing approach. As the number of stages increases, and consequently the size of the problem increases, the proposed method becomes significantly faster than the state-of-the-art decomposition methods. Only in the first instance, the UD model was faster than the CG&S method. This is because of the very small problem size of the first instance. It corresponds to a two-stage problem with the operational stage represented by only one day.

[7] The Progressive Hedging algorithm is considered to be a special case of the more general Alternating Direction Method of Multipliers method (ADMM) [43,44]. The ADMM and the Progressive Hedging algorithm do not guarantee convergence to an optimal solution for mixed-integer integer programming [45,46]. To obtain feasible solutions for mixed-integer programming, additional heuristics are required [46,47].

Table 2
Problem size for different solution methods.

Instance $\|S\|$, $\|\mathcal{N}\|$	# of Variables [Thousand]						# of Constraints [Thousand]					
	UD	CG-ip	CG-ds	CG&S	PH	NBD	UD	CG-ip	CG-ds	CG&S	PH	NBD
2, 4	54.2	18	18	18	18.6	18.3	82.5	27.6	27.6	27.6	28.1	28.2
3, 13	216.2	18	18	18	36.8	18.3	329.2	27.6	27.6	27.6	55.8	28.2
4, 40	702.0	18	18	18	54.3	18.3	1069.3	27.6	27.6	27.6	83.4	28.2
5, 121	2159.6	18	18	18	72.3	18.3	3289.7	27.6	27.6	27.6	111.1	28.2

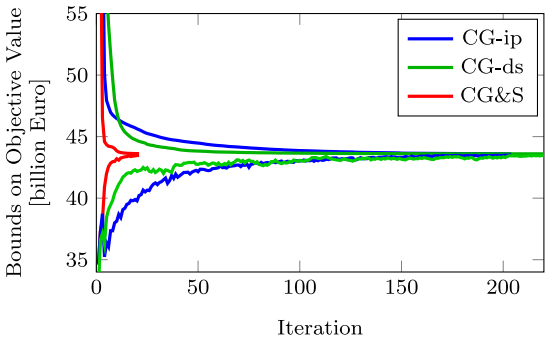

Fig. 5. Convergence of CG with interior point stabilization (CG-ip), CG with dual smoothing (CG-ds) and proposed Column Generation and Sharing (CG&S) method in computational experiments of Section 4.3.

However, the solution time of the UD model increases exponentially with the number of stages. The UD model could not be solved for the largest instance, as expected due to the large size of the problem. This is reflected in the high optimality gap observed in Table 3.

The NBD required a large number of iterations to achieve convergence and showed a pronounced "tailing-off" effect. Only the first instance could be solved within the required gap. Moreover, due to the sequential structure of the NBD, the parallel processing cannot be fully exploited. Therefore, each iteration required more time than for the CG&S approach. The PH algorithm did not allow to solve the problem efficiently. For the first instance, PH required a similar computational effort as the NBD. However, for multi-stage settings, each iteration required a significant effort. Each subproblem is a mixed-integer quadratic problem, which becomes difficult to solve as the number of stages increases.

The proposed Column Generation and Sharing method performs significantly better than the Column Generation with interior point stabilization and Column Generation with dual price smoothing. Compared with the best variant of the CG algorithm, namely CG with dual smoothing, the novel CG&S reduced the solution time from 5.5 h to 46 min in the largest instance. The number of iterations reduced from 222 to 21. A reduction of 86 % in solution time was obtained. Overall, the CG&S method outperforms the CG with interior point stabilization, CG with dual smoothing, NBD, and PH for every instance.

Fig. 5 shows the convergence of the CG with interior point stabilization, CG with dual smoothing, and the Column Generation and Sharing method for the case of five stages. Due to the large number of nodes in the scenario tree and investment options, the CG with interior point stabilization and CG with dual smoothing required a large number of iterations to converge. In the first iterations, there were no sufficient columns in the master problem to obtain useful dual prices. Consequently, the lower bound was not a good approximation, and the "heading-in" effect was observed. Towards the last iterations, the "tailing-off" effect with its slow convergence was observed for the CG with interior point stabilization and CG with dual smoothing. In contrast, with the CS procedure included, the "heading-in" effect was eliminated. The lower bound was significantly improved in the first four iterations. This occurs because the CS procedure is efficient in generating new feasible columns per node of the scenario tree in each iteration, as explained in Section 3.2. These extra columns add more information to the master problem, improving the quality of the dual prices. Also, adding these columns allows to rapidly reduce the upper bound and to decrease the "tailing-off" effect.

In summary, the results show that the proposed Column Generation & Sharing method reduces the solution times considerably compared to state-of-the-art methods. Moreover, it allows to overcome known convergence issues of the standard CG approach.

4.4. Concluding remarks

The results presented in Sections 4.2 and 4.3 demonstrate that the proposed planning framework clearly outperforms the state of the art in terms of solution quality and computational speed. Section 4.2 showed the significant relevance of the improved modeling provided by the proposed integrated planning model. Including UC constraints allows to accurately assess the value of investing in flexibility options. Including long-term uncertainty allows the system planner to anticipate future scenarios. Section 4.3 shows that the proposed solution method reduces solution time by 86 % compared with state-of-the-art-methods. Overall, the proposed planning framework allows to efficiently and effectively model short-term flexibility and long-term uncertainty in a unique planning model.

5. Case study: Value of energy storage systems in low-carbon transition

In this section, the proposed planning framework is applied to a realistic-size case study. The objective of the case study is to quantify the value of ESSs in the transition to a low-carbon power system. The case study considers five stages and 121 nodes in the scenario tree. No reduction was applied to the NREL 118-Bus system described in Section 4.1. The operation is represented by four different weeks of the year as opposed to just one day in Section 4.3. This allows to represent a good number of operating conditions, as well as the seasonality of load and renewable power sources [8]. The approach allows to consider ESSs with hourly to daily energy storage capacity. Two case studies were performed to analyze the value of ESS, a base case with investment in ESSs and a second case without ESSs.

Concerning the computational performance, the case with ESSs was solved in eight days with the proposed framework, a 1.5 % gap was obtained, and 54 iterations were needed. By this same time, the NBD method had performed 21 iterations obtaining a gap higher than 100 %, and the PH method could not complete the first iteration. These results confirm the capability of the proposed planning framework in overcoming the computational burden when solving realistic-size problems.

Fig. 6 shows the expected installed generation capacity for both studied cases. Stage 1 is omitted as it represents the initial state of the system. In both cases, investment in new generation is driven by the CO_2 emission limit imposed by (24). Therefore, an increasing capacity of renewable power is installed at each stage to achieve the low-carbon transformation. The decreasing investment costs of solar generation and energy storage systems, and the increasing cost of fossil fuels also contribute to the higher integration of renewable sources. The case with ESSs allows for a higher installed capacity of renewables up to stage 4, compared with the case without ESSs. In stage 5, the opposite is

Table 3
Computational performance.

Instance $\|S\|$, $\|\mathcal{N}\|$	Solution time [min]						# of Iterations					Gap [%]		
	UD	CG-ip	CG-ds	CG&S	PH	NBD	CG-ip	CG-ds	CG&S	PH	NBD	UD	PH	NBD
2, 4	0.16	15.5	3.8	1.7	8.0	8.3	152	80	10	9	112	<0.5	<0.5	<0.5
3, 13	14.0	26.9	23.0	5.2	543.2	1440	175	168	23	35	1699	<0.5	2.4	1.5
4, 40	881.4	202.6	93.6	16.8	1440	1440	168	203	27	33	480	<0.5	2.1	1.8
5, 121	1440	389.4	330.4	46.5	1440	1440	204	222	21	10	336	98.6	2.7	1.6

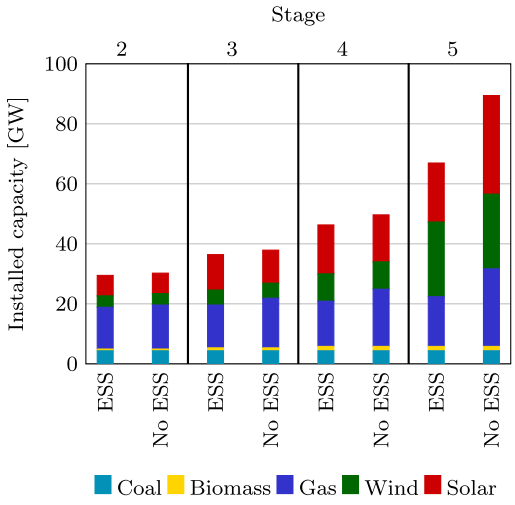

Fig. 6. Expected installed generation capacities in case study of Section 5.

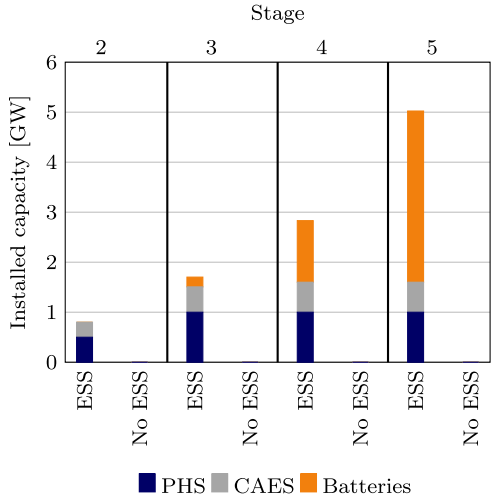

Fig. 7. Expected installed power capacity of energy storage systems in case study of Section 5.

observed. In the case with ESSs, the amount of solar power installed is reduced by 13 GW compared to the case without ESSs. This is due to the flexibility provided by ESSs, where available RES can be integrated more efficiently into the system. In the case with no ESSs, a higher capacity of gas-fired units has to be installed in each stage as back-up generation to mitigate the variability of the renewable sources and to provide ancillary services.

Fig. 7 shows the expected installed power capacity of energy storage systems. In stage 2, corresponding to the period 2021 to 2025, a relatively small power capacity of pumped hydro storage and CAES is installed in the system. This capacity amounts to less than 3 % of the total installed generation capacity. In this stage, pumped hydro storage and CAES are preferred over batteries due to their lower investment costs and longer lifetime. These technologies, especially pumped hydro storage, are in a more mature state [49]. In stage 3, the installed power capacity of PHS and CAES is increased to support the rising share of renewable generation. In stage 4, corresponding to the period 2031 to 2035, the stringent CO_2 emission limit leads to a share of 54.4% of renewable power generation. A significant amount of flexibility is required to efficiently integrate this high penetration of renewable sources. Therefore, batteries become a cost-efficient alternative to provide flexibility to the power system. The installed power capacity of ESSs in the final stage represents 7 % of the total installed capacity in the system.

Fig. 8 shows the expected yearly electricity generation for both cases. ESSs allowed for a more efficient penetration of renewable energy, with a significant reduction of renewable energy curtailment. Energy storage systems also allowed for a lower utilization of gas. Fig. 9 shows the expected yearly CO_2 emissions by fuel type. In both cases, the increasing penetration of renewable generation allows reducing the CO_2 emissions to comply with the emission limit imposed by (24). Coal-fired generation is replaced by renewable generation. Thus, CO_2 emissions from coal-fired generation decrease rapidly. The CO_2 emissions from gas-fired generation decrease more slowly. The case with ESSs allowed for a higher utilization of coal-fired units and a lower

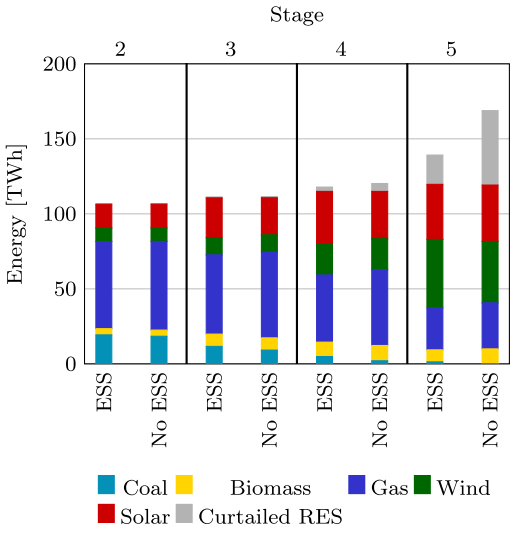

Fig. 8. Expected yearly electricity generation in case study of Section 5.

utilization of the more expensive gas-fired generation. Therefore, CO_2 emissions originating from coal-fired generation are higher in the case with ESSs compared to the case without ESSs.

Table 4 shows the system costs for the case with and without ESSs. The case without ESSs results in higher investment costs because of the higher investment in gas-fired units and solar power. Fuel and start-up costs are also higher in the case without ESSs because of the higher renewable curtailment and the higher utilization of gas-fired generation. Load shedding occurs in the case without ESSs, increasing the system cost significantly. Overall, in the case without ESSs the expected system cost was 18 % higher than in the case with ESSs.

Table 4
Expected system costs in case study of Section 5.

Case	Investment cost [billion Euro]	Fuel and start-up cost [billion Euro]	Load shedding cost [billion Euro]	Total system cost [billion Euro]
With ESSs	13.11	36.46	0.01	49.58
Without ESSs	14.75	38.74	5.13	58.62

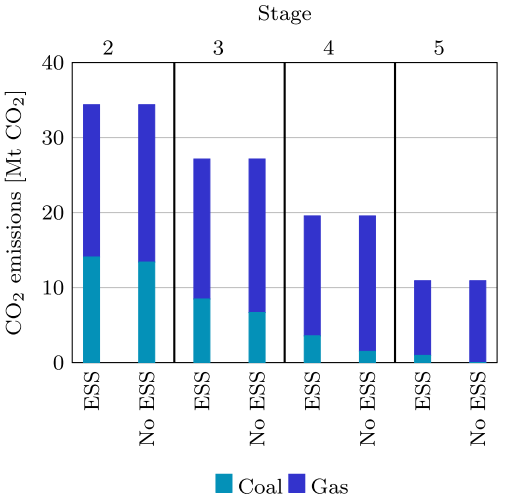

Fig. 9. Expected yearly CO_2 emissions by fuel type in case study of Section 5.

6. Conclusions

A comprehensive long-term power system planning framework that integrates short-term flexibility and long-term uncertainty was proposed. The planning model considers investments in generation, transmission, and energy storage. A general model for energy storage systems was developed to represent the characteristics of the different technologies. The model allows an accurate representation of the short-term flexibility requirements through the unit commitment constraints. Such a comprehensive model benefits from an efficient solution method. Therefore, a distributed computing framework based on the novel Column Generation and Sharing algorithm was proposed. The added value offered by the proposed method was analyzed on case studies based on the NREL 118-bus system. Results show that ignoring unit commitment constraints leads to suboptimal solutions with insufficient investment in flexibility options and significantly higher cost. Neglecting long-term uncertainty results in a system that is not capable to cope with the different long-term scenarios. The computational study clearly showed the improvement in solution speed obtained by the proposed framework compared to state-of-the-art methods. It reduces the solution time by 86 % and allows to solve realistic-size planning instances that cannot be solved by existing methods. Results also confirm the value of energy storage systems in the long-term transition to a low-carbon power system. As such, the combination of accuracy and superior computational speed makes the proposed framework a compelling candidate for planning the transition to zero-carbon power systems.

CRediT authorship contribution statement

Angela Flores-Quiroz: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Kai Strunz:** Conceptualization, Methodology, Validation, Writing - review & editing, Visualization, Supervision.

Acknowledgments

Support was given by CONICYT-DAAD, Chile/DOCTORADO/2014 and by project EchtEWende (0325814 A) of BMWi, Germany. This research was partially supported by the supercomputing infrastructure of the NLHPC, Chile (ECM-02).

Appendix A. Dantzig–Wolfe decomposition

To facilitate a decomposition, the matrix form (A.1) of the multi-stage stochastic problem (48) is introduced first. Constraint (A.1b) summarizes the non-anticipativity constraints (2)–(4). These are the complicating constraints that couple the investment decisions made at the predecessor nodes of n with the total installed units in node n. Eq. (A.1c) represents the constraints that relate operational decisions in node n with the total installed units in node n, namely (6), (8), (10), (34)–(35), and (38)–(47). Constraint (A.1d) summarizes the maximum investment limit, right-hand side of (2)–(4). Constraint (A.1e) summarizes the operational constraints that do not explicitly depend on the total installed units, namely (5), (9), (21)–(33), (36), and (37). Constraint (A.1f) states the integrality of investment decisions. Constraints (A.1g)–(A.1i) summarize the constraints on the investment decisions of existing assets, namely constraints (12)–(20). Problem (A.1) has a block diagonal structure, where a set of independent subproblems is linked by a small number of complicating constraints [31].

$$\min \quad \sum_{n\in\mathcal{N}} \phi_n \left(\boldsymbol{c}_n^{\text{inv}\top} \boldsymbol{I}_n + \boldsymbol{c}_n^{\text{op}\top} \boldsymbol{X}_n \right) \tag{A.1a}$$

$$\text{s.t.:} \quad \boldsymbol{Z}_n \le \sum_{h\in\mathcal{P}_n} \boldsymbol{I}_h \qquad \forall\, n \tag{A.1b}$$

$$\boldsymbol{A}_n \boldsymbol{X}_n \le \boldsymbol{Z}_n \qquad \forall\, n \tag{A.1c}$$

$$\boldsymbol{Z}_n \le \boldsymbol{Z}_n^{\max} \qquad \forall\, n \tag{A.1d}$$

$$\boldsymbol{X}_n \in \mathcal{X}_n \qquad \forall\, n \tag{A.1e}$$

$$\boldsymbol{Z}_n, \boldsymbol{I}_n \in \mathbb{Z}_+^{|G|+|\mathcal{E}|+|L|} \qquad \forall\, n \tag{A.1f}$$

$$\text{diag}(\boldsymbol{U}^0)\boldsymbol{Z}_n = \boldsymbol{Z}^0 \qquad \forall\, n \tag{A.1g}$$

$$\text{diag}(\boldsymbol{U}^0)\boldsymbol{I}_1 = \boldsymbol{Z}^0 \tag{A.1h}$$

$$\text{diag}(\boldsymbol{U}^0)\boldsymbol{I}_n = \boldsymbol{0} \qquad \forall\, n > 1 \tag{A.1i}$$

To apply the Dantzig–Wolfe decomposition, the problem is reformulated using the discretization approach [22,31]. To do so, the feasible region of total installed units in node n, which includes generators, ESSs, and lines, is defined as:

$$\mathcal{Z}_n = \{ \boldsymbol{Z}_n \in \mathbb{Z}_+^{|G|+|\mathcal{E}|+|L|} \mid \exists \boldsymbol{X}_n \in \mathcal{X}_n,\ \boldsymbol{A}_n\boldsymbol{X}_n \le \boldsymbol{Z}_n \le \boldsymbol{Z}_n^{\max},\ \text{diag}(\boldsymbol{U}^0)\boldsymbol{Z}_n = \boldsymbol{Z}^0 \}$$

where $\mathcal{Z}_n$ is a bounded integer polyhedron. Therefore, any point in $\mathcal{Z}_n$ can be expressed as a combination of a finite number of integer points, $\{\hat{\boldsymbol{Z}}_n^j\}_{j\in K_n}$, in $\mathcal{Z}_n$ [31] such that

$$\boldsymbol{Z}_n = \sum_{j\in K_n} \lambda_n^j \hat{\boldsymbol{Z}}_n^j, \quad \sum_{j\in K_n} \lambda_n^j = 1, \lambda_n^j \in \{0,1\}. \tag{A.2}$$

For each feasible vector of total installed units in node n, $\hat{\boldsymbol{Z}}_n^j$, there exists at least one associated optimal operational plan, $\hat{\boldsymbol{X}}_n^j$, therefore $\boldsymbol{X}_n$ can be represented as:

$$\boldsymbol{X}_n = \sum_{j\in K_n} \lambda_n^j \hat{\boldsymbol{X}}_n^j. \tag{A.3}$$

The master problem of the Dantzig–Wolfe decomposition (A.4) is obtained by substituting $\boldsymbol{Z}_n$ and $\boldsymbol{X}_n$ in (A.1) by (A.2) and (A.3), respectively. Constraint (A.4b) is equivalent to the non-anticipativity constraints (A.1b). Constraints (A.4c) and (A.4d) ensure that only one vector of total installed units, and consequently only one operational plan, is selected for each node in the scenario tree. The associated dual prices of constraints (A.4b) and (A.4c) are $\boldsymbol{\pi}_n$ and μ_n, respectively.

$$\min \quad \sum_{n\in\mathcal{N}} \phi_n(\boldsymbol{c}_n^{\mathrm{inv}\top}\boldsymbol{I}_n + \sum_{j\in K_n} \lambda_n^j \boldsymbol{c}_n^{\mathrm{op}\top}\hat{\boldsymbol{X}}_n^j) \tag{A.4a}$$

$$\text{s.t.:} \quad \sum_{j\in K_n} \lambda_n^j \hat{\boldsymbol{Z}}_n^j \le \sum_{h\in\mathcal{P}_n} \boldsymbol{I}_h \qquad \forall n \quad (\boldsymbol{\pi}_n) \tag{A.4b}$$

$$\sum_{j\in K_n} \lambda_n^j = 1 \qquad \forall n \quad (\mu_n) \tag{A.4c}$$

$$\lambda_n^j \in \{0,1\} \qquad \forall n, j \tag{A.4d}$$

$$\boldsymbol{I}_n \in \mathbb{Z}_+^{|G|+|\mathcal{E}|+|L|} \qquad \forall n \tag{A.4e}$$

$$\mathrm{diag}(\boldsymbol{U}^0)\boldsymbol{I}_1 = \boldsymbol{Z}^0 \tag{A.4f}$$

$$\mathrm{diag}(\boldsymbol{U}^0)\boldsymbol{I}_n = \boldsymbol{0} \qquad \forall n > 1 \tag{A.4g}$$

Problem (A.4) may be solved by the CG approach [31]. The CG approach allows to obtain columns $\{\hat{\boldsymbol{Z}}_n^j, \hat{\boldsymbol{X}}_n^j\}$ by solving subproblem SP_n, defined by (A.5), for each node. The objective value of the subproblem, denoted by z_n^{sp}, is defined as the reduced cost of column $\{\boldsymbol{Z}_n, \boldsymbol{X}_n\}$.

$$(SP_n) \quad z_n^{\mathrm{sp}} = \min \quad \phi_n \boldsymbol{c}_n^{\mathrm{op}\top}\boldsymbol{X}_n - \boldsymbol{\pi}_n^\top \boldsymbol{Z}_n - \mu_n \tag{A.5a}$$

$$\text{s.t.:} \quad \boldsymbol{X}_n \in \mathcal{X}_n \tag{A.5b}$$

$$\boldsymbol{A}_n\boldsymbol{X}_n \le \boldsymbol{Z}_n \tag{A.5c}$$

$$\mathrm{diag}(\boldsymbol{U}^0)\boldsymbol{Z}_n = \boldsymbol{Z}^0 \tag{A.5d}$$

$$\boldsymbol{Z}_n \le \boldsymbol{Z}_n^{\max}\,,\ \boldsymbol{Z}_n \in \mathbb{Z}_+^{|G|+|\mathcal{E}|+|L|} \tag{A.5e}$$

Appendix B. Scenario generation method

To generate the scenario tree, the methodology proposed in [39,40] is used. Demand and gas price are modeled as two correlated geometric Brownian motion processes. Then, the scenario tree is obtained using the moment matching method [50] with three branches per node. The aim is to find vector $\boldsymbol{y}_n^\top = (y_{1n}, y_{2n}, \ldots, y_{Jn})$ and a scalar $\bar{\phi}_n$ for each node n of the scenario tree. Vector $\boldsymbol{y_n}$ describes the value of the uncertain parameters at node n. The conditional probability associated with node n is defined by $\bar{\phi}_n$. The objective of the moment matching method is to obtain $\boldsymbol{y_n}$ and $\bar{\phi}_n$ such that the statistical properties of the original stochastic process match the statistical properties of the approximated distribution [50]. For multi-stage scenario trees, the moment matching method generates the scenario tree recursively, starting from the root node until the leaf nodes are reached [39,40,50]. In each node n, the following optimization problem is solved:

$$\min_{\boldsymbol{y}_{\tilde{n}},\bar{\phi}_{\tilde{n}}} \quad \sum_{i\in\mathcal{Q}} \varphi_i (f_i(\{\boldsymbol{y}_{\tilde{n}}\}_{\tilde{n}\in D_n}, \{\bar{\phi}_{\tilde{n}}\}_{\tilde{n}\in D_n}) - Q_i^{\mathrm{VAL}}(\boldsymbol{y}_n))^2 \tag{B.1a}$$

$$\text{s.t.:} \quad \sum_{\tilde{n}\in D_n} \bar{\phi}_{\tilde{n}} = 1 \tag{B.1b}$$

$$\bar{\phi}_{\tilde{n}} \ge 0 \qquad \forall \tilde{n} \in D_n \tag{B.1c}$$

In (B.1), $\mathcal{Q}$ is the set of statistical properties. The value of the ith statistical property computed for the underlying distribution of the stochastic process is described by $Q_i^{\mathrm{VAL}}(\boldsymbol{y}_n)$. The value of the ith statistical property computed for $\{\boldsymbol{y}_{\tilde{n}}\}_{\tilde{n}\in D_n}$ and $\{\bar{\phi}_{\tilde{n}}\}_{\tilde{n}\in D_n}$ is denoted by $f_i(\{\boldsymbol{y}_{\tilde{n}}\}_{\tilde{n}\in D_n}, \{\bar{\phi}_{\tilde{n}}\}_{\tilde{n}\in D_n})$. The weights φ_i specify the relative importance of the ith statistical property. Set D_n describe the set of children nodes of node n. Following [39,40], the selected statistical properties are the mean, the variance, the skewness of each random variable, and the correlation of the random variables.

Appendix C. Stabilization methods for column generation approach

As described in Section 3.2, the standard Column Generation algorithm presents convergence issues [31], such as the "heading-in" and "tailing-off" effects. Stabilization techniques are usually used to overcome these issues [14,22,41,42,51]. These techniques aim to accelerate the convergence of the Column Generation by controlling the behavior of the dual prices obtained by the master problem. In [41], the interior point stabilization is proposed. It aims to improve the convergence by using dual prices that are interior point solutions to the master problem. This technique can be implemented by solving the master problem with the interior point method instead of using Simplex-based methods [14,22]. Dual smoothing techniques [42,51] improve the convergence of the CG method by correcting the dual prices to obtain values closer to the stability center. In [51], dual smoothing is achieved by using the static Wentges smoothing. In [42], the auto-adaptive dual smoothing is proposed to improve the Wentges smoothing by eliminating the need for application-specific parameter tuning.

In contrast to stabilization methods, the proposed Column Generation and Sharing method overcomes convergence issues from a different perspective. As described in Section 3.2, it aims to provide more information to the master problem in each iteration. This is achieved by the proposed Column Sharing procedure, which efficiently generates relevant columns for the master problem by exploiting the scenario tree structure. Furthermore, the discussed state-of-the-art stabilization techniques can also be used within the proposed Column Generation and Sharing method.

Appendix D. Input data for validation and case study

In the following, the input data for the validation of performance of Section 4 and the case study of Section 5 is given. Table D.5 presents the expected demand and fuel prices in each stage of the planning horizon. The demand at the initial stage is obtained from [34]. The expected demand growth rate is based on [40]. Expected fuel prices are given as in [38]. The expected demand and fuel prices are used as input for the scenario tree generation. Table D.6 presents the CO_2 emission limit. Energy storage system parameters are given in Table D.7, based on [36,37]. Generator parameters are given in Table D.8, based on [35]. Fig. D.10 shows the investment costs for the different technologies, based on [35,36]. Upward primary reserve requirement was equal to 4 % of the load. Secondary reserve requirement was defined as 3 % of the load plus 5 % of the variable generation output.

Table D.5
Expected demand and fuel prices.

Stage	Demand [TWh]	Coal price [Euro/MWh$_{th}$]	Natural gas price [Euro/MWh$_{th}$]
1	102.4	7.7	26.3
2	106.5	7.9	30.4
3	110.7	9.3	32.2
4	115.0	11.2	33.9
5	119.4	12.9	35.9

Table D.6
CO_2 emission limit.

Stage	Year	CO_2 emission limit [Mt CO_2]
1	2016	54.5
2	2025	34.4
3	2030	27.1
4	2035	19.6
5	2040	10.9

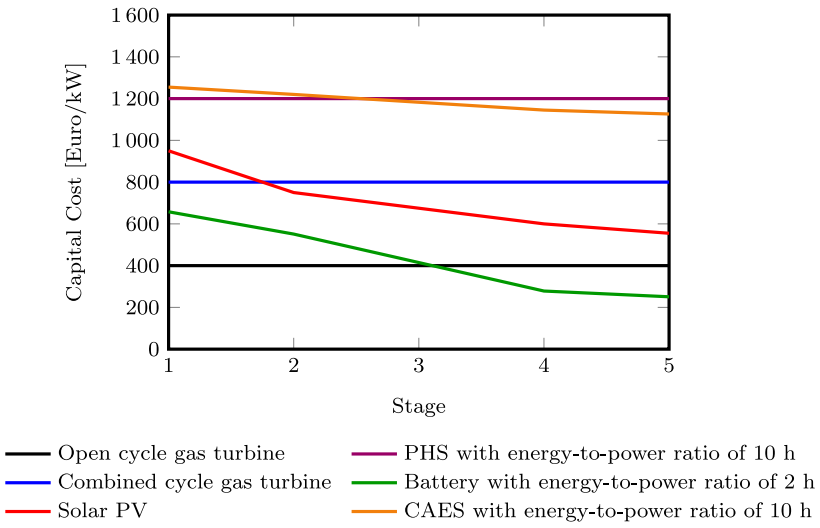

Fig. D.10. Investment costs of different technologies.

Table D.7
Technical parameters of energy storage systems.

Technology	Energy-to-power ratio [h]	Charge efficiency [%]	Discharge efficiency [%]	Heat rate [MWh_{th}/MWh]	Emission factor [t CO_2/MWh]	Ramp rate as share of max. output power [%/min]	Lifetime [years]
PHS	10	0.85	0.85	0	0	50	80
CAES	10	1	1.2[a]	1.23	0.23	20	30
Battery	2	0.95	0.95	0	0	100	8

[a]CAES is modeled with an efficiency greater than one because of the use of natural gas in discharging mode [37,52].

Table D.8
Technical parameters of generators.

Technology	Efficiency [%]	Emission factor [t CO_2/MWh]	Min. output as share of max. output power [%]	Ramp rate as share of max. output power [%/min]	Minimum time on [h]	Minimum time off [h]	Lifetime [years]
Biomass	0.45	0	0	20	1	1	25
Combined cycle gas turbine	0.57	0.349	40	8	4	4	40
Open cycle gas turbine	0.45	0.488	20	15	1	1	50
Coal	0.43	0.728	40	4	24	24	45
Solar PV	–	–	–	–	–	–	25
Wind	–	–	–	–	–	–	25

References

[1] Lannoye E, Flynn D, O'Malley M. Transmission, variable generation, and power system flexibility. IEEE Trans Power Syst 2015;30(1):57–66. http://dx.doi.org/10.1109/TPWRS.2014.2321793.

[2] Heggarty T, Bourmaud J-Y, Girard R, Kariniotakis G. Multi-temporal assessment of power system flexibility requirement. Appl Energy 2019;238:1327–36. http://dx.doi.org/10.1016/j.apenergy.2019.01.198.

[3] Heggarty T, Bourmaud J-Y, Girard R, Kariniotakis G. Quantifying power system flexibility provision. Appl Energy 2020;279:115852. http://dx.doi.org/10.1016/j.apenergy.2020.115852.

[4] Maloney P, Chitkara P, McCalley J, Hobbs B, Clack C, Ortega-Vazquez M, et al. Research to develop the next generation of electric power capacity expansion tools: What would address the needs of planners?. Int J Electr Power Energy Syst 2020;121:106089. http://dx.doi.org/10.1016/j.ijepes.2020.106089.

[5] Palmintier BS, Webster MD. Impact of operational flexibility on electricity generation planning with renewable and carbon targets. IEEE Trans Sustain Energy 2016;7(2):672–84. http://dx.doi.org/10.1109/TSTE.2015.2498640.

[6] de Sisternes FJ, Jenkins JD, Botterud A. The value of energy storage in decarbonizing the electricity sector. Appl Energy 2016;175:368–79. http://dx.doi.org/10.1016/j.apenergy.2016.05.014.

[7] Moreno R, Street A, Arroyo JM, Mancarella P. Planning low-carbon electricity systems under uncertainty considering operational flexibility and smart grid technologies. Phil Trans R Soc A 2017;375(2100). http://dx.doi.org/10.1098/rsta.2016.0305.

[8] Ma J, Silva V, Belhomme R, Kirschen D, Ochoa L. Evaluating and planning flexibility in sustainable power systems. IEEE Trans Sustain Energy 2013;4(1):200–9. http://dx.doi.org/10.1109/TSTE.2012.2212471.

[9] Welsch M, Deane P, Howells M, Ó Gallachóir B, Rogan F, Bazilian M, et al. Incorporating flexibility requirements into long-term energy system models – A case study on high levels of renewable electricity penetration in Ireland. Appl Energy 2014;135:600–15. http://dx.doi.org/10.1016/j.apenergy.2014.08.072.

[10] Hua B, Baldick R, Wang J. Representing operational flexibility in generation expansion planning through convex relaxation of unit commitment. IEEE Trans Power Syst 2018;33(2):2272–81. http://dx.doi.org/10.1109/TPWRS.2017.2735026.

[11] Du E, Zhang N, Kang C, Xia Q. A high-efficiency network-constrained clustered unit commitment model for power system planning studies. IEEE Trans Power Syst 2019;34(4):2498–508. http://dx.doi.org/10.1109/TPWRS.2018.2881512.

[12] Manríquez F, Sauma E, Aguado J, de la Torre S, Contreras J. The impact of electric vehicle charging schemes in power system expansion planning. Appl Energy 2020;262:114527. http://dx.doi.org/10.1016/j.apenergy.2020.114527.

[13] Koltsaklis NE, Georgiadis MC. A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. Appl Energy 2015;158:310–31. http://dx.doi.org/10.1016/j.apenergy.2015.08.054.

[14] 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. Electr Power Syst Res 2016;136:232–41. http://dx.doi.org/10.1016/j.epsr.2016.02.011.

[15] Quiroga D, Sauma E, Pozo D. Power system expansion planning under global and local emission mitigation policies. Appl Energy 2019;239:1250–64. http://dx.doi.org/10.1016/j.apenergy.2019.02.001.

[16] Park H, Baldick R. Multi-year stochastic generation capacity expansion planning under environmental energy policy. Appl Energy 2016;183:737–45. http://dx.doi.org/10.1016/j.apenergy.2016.08.164.

[17] Liu Y, Sioshansi R, Conejo AJ. Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. IEEE Trans Power Syst 2018;33(1):781–91. http://dx.doi.org/10.1109/TPWRS.2017.2694612.
[18] Falugi P, Konstantelos I, Strbac G. Planning with multiple transmission and storage investment options under uncertainty: A nested decomposition approach. IEEE Trans Power Syst 2018;33(4):3559–72. http://dx.doi.org/10.1109/TPWRS.2017.2774367.
[19] Lara CL, Siirola JD, Grossmann IE. Electric power infrastructure planning under uncertainty: stochastic dual dynamic integer programming (SDDiP) and parallelization scheme. Optim Eng 2019. http://dx.doi.org/10.1007/s11081-019-09471-0.
[20] Lohmann T, Hering AS, Rebennack S. Spatio-temporal hydro forecasting of multireservoir inflows for hydro-thermal scheduling. European J Oper Res 2016;255(1):243–58. http://dx.doi.org/10.1016/j.ejor.2016.05.011.
[21] Dentcheva D, Römisch W. Duality gaps in nonconvex stochastic optimization. Math Program 2004;101(3):515–35. http://dx.doi.org/10.1007/s10107-003-0496-1.
[22] Singh KJ, Philpott AB, Wood RK. Dantzig-Wolfe decomposition for solving multistage stochastic capacity-planning problems. Oper Res 2009;57(5):1271–86. http://dx.doi.org/10.1287/opre.1080.0678.
[23] Romero R, Monticelli A, Garcia A, Haffner S. Test systems and mathematical models for transmission network expansion planning. IEE Proc Gener Transm Distrib 2002;149(1):27–36. http://dx.doi.org/10.1049/ip-gtd:20020026.
[24] May GJ, Davidson A, Monahov B. Lead batteries for utility energy storage: A review. J Energy Storage 2018;15:145–57. http://dx.doi.org/10.1016/j.est.2017.11.008.
[25] Perez A, Moreno R, Moreira R, Orchard M, Strbac G. Effect of battery degradation on multi-service portfolios of energy storage. IEEE Trans Sustain Energy 2016;7(4):1718–29. http://dx.doi.org/10.1109/TSTE.2016.2589943.
[26] Holttinen H, Milligan M, Ela E, Menemenlis N, Dobschinski J, Rawn B, et al. Methodologies to determine operating reserves due to increased wind power. IEEE Trans Sustain Energy 2012;3(4):713–23. http://dx.doi.org/10.1109/TSTE.2012.2208207.
[27] van Stiphout A, De Vos K, Deconinck G. The impact of operating reserves on investment planning of renewable power systems. IEEE Trans Power Syst 2017;32(1):378–88. http://dx.doi.org/10.1109/TPWRS.2016.2565058.
[28] Bylling HC, Pineda S, Boomsma TK. The impact of short-term variability and uncertainty on long-term power planning. Ann Oper Res 2020;284(1):199–223. http://dx.doi.org/10.1007/s10479-018-3097-3.
[29] Rebours YG, Kirschen DS, Trotignon M, Rossignol S. A survey of frequency and voltage control ancillary services—Part I: Technical features. IEEE Trans Power Syst 2007;22(1):350–7. http://dx.doi.org/10.1109/TPWRS.2006.888963.
[30] Duan C, Jiang L, Fang W, Liu J, Liu S. Data-driven distributionally robust energy-reserve-storage dispatch. IEEE Trans Ind Inf 2018;14(7):2826–36. http://dx.doi.org/10.1109/TII.2017.2771355.
[31] Vanderbeck F. Implementing mixed integer column generation. In: Column generation. Boston, MA: Springer US; 2005, p. 331–58. http://dx.doi.org/10.1007/0-387-25486-2_12.
[32] Desrosiers J, Lübbecke ME. A primer in column generation. In: Column generation. Boston, MA: Springer US; 2005, p. 1–32. http://dx.doi.org/10.1007/0-387-25486-2_1.
[33] Vigerske S. Decomposition in multistage stochastic programming and a constraint integer programming approach to mixed-integer nonlinear programming (Ph.D. thesis), Humboldt-Univ. Berlin; 2012.
[34] Peña I, Martinez-Anido CB, Hodge B. An extended IEEE 118-Bus test system with high renewable penetration. IEEE Trans Power Syst 2018;33(1):281–9. http://dx.doi.org/10.1109/TPWRS.2017.2695963.
[35] Schröder A, Kunz F, Meiss J, Mendelevitch R, von Hirschhausen C. Current and prospective costs of electricity generation until 2050. Data Documentation 68, DIW Berlin; 2013.
[36] Gerbaulet C, von Hirschhausen C, Kemfert C, Lorenz C, Oei P-Y. European electricity sector decarbonization under different levels of foresight. Renew Energy 2019;141:973–87. http://dx.doi.org/10.1016/j.renene.2019.02.099.
[37] Cleary B, Duffy A, O'Connor A, Conlon M, Fthenakis V. Assessing the economic benefits of compressed air energy storage for mitigating wind curtailment. IEEE Trans Sustain Energy 2015;6(3):1021–8. http://dx.doi.org/10.1109/TSTE.2014.2376698.
[38] Capros P, De Vita A, Tasios N, Siskos P, Kannavou M, Petropoulos A, et al. EU reference scenario 2016: energy, transport and GHG emissions – trends to 2050. European Comission; 2016.
[39] Jin S, Ryan SM, Watson J-P, Woodruff DL. Modeling and solving a large-scale generation expansion planning problem under uncertainty. Energy Syst 2011;2(3):209–42. http://dx.doi.org/10.1007/s12667-011-0042-9.
[40] Feng Y, Ryan SM. Scenario construction and reduction applied to stochastic power generation expansion planning. Comput Oper Res 2013;40(1):9–23. http://dx.doi.org/10.1016/j.cor.2012.05.005.
[41] Rousseau L-M, Gendreau M, Feillet D. Interior point stabilization for column generation. Oper Res Lett 2007;35(5):660–8. http://dx.doi.org/10.1016/j.orl.2006.11.004.
[42] Pessoa A, Sadykov R, Uchoa E, Vanderbeck F. Automation and combination of linear-programming based stabilization techniques in column generation. INFORMS J Comput 2018;30(2):339–60. http://dx.doi.org/10.1287/ijoc.2017.0784.
[43] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn 2011;3(1):1–122. http://dx.doi.org/10.1561/2200000016.
[44] Boland N, Christiansen J, Dandurand B, Eberhard A, Linderoth J, Luedtke J, et al. Combining progressive hedging with a Frank–Wolfe method to compute Lagrangian dual bounds in stochastic mixed-integer programming. SIAM J Optim 2018;28(2):1312–36. http://dx.doi.org/10.1137/16M1076290.
[45] Takapoui R, Moehle N, Boyd S, Bemporad A. A simple effective heuristic for embedded mixed-integer quadratic programming. Internat J Control 2020;93(1):2–12. http://dx.doi.org/10.1080/00207179.2017.1316016.
[46] Watson J-P, Woodruff DL. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. Comput Manage Sci 2011;8(4):355–70. http://dx.doi.org/10.1007/s10287-010-0125-4.
[47] Feizollahi MJ, Costley M, Ahmed S, Grijalva S. Large-scale decentralized unit commitment. Int J Electr Power Energy Syst 2015;73:97–106. http://dx.doi.org/10.1016/j.ijepes.2015.04.009.
[48] Munoz FD, Watson J-P. A scalable solution framework for stochastic transmission and generation planning problems. Comput Manage Sci 2015;12(4):491–518. http://dx.doi.org/10.1007/s10287-015-0229-y.
[49] Zafirakis D, Chalvatzis KJ, Baiocchi G, Daskalakis G. The value of arbitrage for energy storage: Evidence from European electricity markets. Appl Energy 2016;184:971–86. http://dx.doi.org/10.1016/j.apenergy.2016.05.047.
[50] Hoyland K, Wallace SW. Generating scenario trees for multistage decision problems. Manage Sci 2001;47(2):295–307. http://dx.doi.org/10.1287/mnsc.47.2.295.9834.
[51] Kowalczyk D, Leus R. A branch-and-price algorithm for parallel machine scheduling using ZDDs and generic branching. INFORMS J Comput 2018;30(4):768–82. http://dx.doi.org/10.1287/ijoc.2018.0809.
[52] Drury E, Denholm P, Sioshansi R. The value of compressed air energy storage in energy and reserve markets. Energy 2011;36(8):4959–73. http://dx.doi.org/10.1016/j.energy.2011.05.041.