Generation and Transmission Expansion Planning Towards a 100% Renewable Future

Shengfei Yin, *Student Member, IEEE*, and Jianhui Wang, *Senior Member, IEEE*

***Abstract*—This paper proposes a novel modeling framework and decomposition-based solution strategy combining stochastic programming (SP) and robust optimization (RO) to deal with multiplex uncertainties in coordinated mid- and long-term power system planning. The problem is formulated as a multi-year generation and transmission planning problem from an independent system operator (ISO)'s perspective to minimize both expansion and operational costs under binary and continuous uncertainties, i.e., system component contingency and load/generation variation. N-k contingencies are captured in RO using the reformulated contingency criteria, while the load/generation uncertainty is considered in SP embedded with RO using operating scenarios generated from the historical data with spatiotemporal correlations. The original hybrid model is highly intractable, but the intractability can be relieved by the proposed decomposition strategy based on the column-and-constraint generation and L-shaped algorithms. We apply our model to perform long-term system planning under extremely high renewable penetration and investigate the case of 100% renewables in long-term planning. Numerical experiments on multi-scale test systems verify the efficacy of the proposed approach.**

***Index Terms*—Coordinated planning, discrete and continuous uncertainties, stochastic programming, robust optimization, decomposition, 100% renewable penetration.**

NOMENCLATURE

SETS:

$G\,/\,R\,/\,L$	Conventional generator / renewable generator / transmission line
$S\,/\,V$	Sending/receiving bus of transmission lines
C	Bus mapping of conventional generators, renewable generators, and demand
$X^G\,/\,X^R$	Candidate conventional / renewable generator
X^L	Candidate transmission line
T	Planning horizon (years)
H	Operation horizon (hours)
M	Planning / operation time index mapping
N	Control mapping between units and investable years

Manuscript received October 9, 2019; revised March 3, 2020, June 5, 2020, August 18, 2020, and September 10, 2020; accepted October 19, 2020. This work supported by the Department of Energy under Award Number(s) DE-OE0000842. *(Corresponding author: Jianhui Wang.)*

The authors are with the Lyle School of Engineering, Southern Methodist University, Dallas, TX 75275-0221 USA (e-mail: gyin@smu.edu; jianhui@smu.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2020.3033487

INDICES:

$g\,/\,r\,/\,\ell$	Conventional generator / renewable generator / transmission line
d	Demand
n	Bus
ω	Scenario
k	Iteration counter for C&CG algorithm
o	Iteration counter for L-shaped algorithm
$t\,/\,h$	Planning / operating time index. $(t,h)\in M$

PARAMETERS:

$IC^G_{g,t}/IC^R_{r,t}$	Investment cost for gth conventional / rth renewable generator in tth year [$]
$IC^L_{\ell,t}$	Investment cost for ℓth transmission line in tth year [$]
$OC_{g,h}$	Operating cost for gth generator in hth hour [$/MWh]
$PC_{d,h}$	Load-shedding cost for unserved load in dth demand in hth hour [$/MWh]
$B^G\,/\,B^L$	Investment budget for generators / transmission lines [$]
K_t	Contingency criterion for system units in tth year
A_ℓ	Line reactance of ℓth transmission line under base MVA [p.u.]
FL_ℓ	Flow capacity of ℓth transmission line [MW]
$\overline{PL}^G_g$	Capacity limit of gth conventional generator [MW]
$\underline{\Delta}_n\,/\,\overline{\Delta}_n$	Angle limit of phase angle at nth bus [rad]

UNCERTAIN PARAMETERS:

$\overline{PL}^R_{r,h}(\omega)$	Available active power of rth renewable generator in hth hour [MW].
$P_{d,h}(\omega)$	Active demand of dth load in hth hour [MW].

VARIABLES:

$x^G_{g,t}/x^R_{r,t}$	Binary expansion decision for gth conventional / rth renewable generator in tth year. $x=1$ means built; $x=0$ otherwise.
$x^L_{\ell,t}$	Binary expansion decision for ℓth transmission line. $x=1$ means built; $x=0$ otherwise.
$y^G_{g,t}/y^R_{r,t}$	Binary availability indicator for gth conventional / rth renewable generator in tth year. $y=1$ means available; $y=0$ otherwise.
$y^L_{\ell,t}$	Binary availability indicator for ℓth transmission line. $y=1$ means available; $y=0$ otherwise.

$a^G_{g,t}$ / $a^R_{r,t}$ Binary outage indicator for gth / rth generator. $a = 1$ means outage; $a = 0$ otherwise.

$a^L_{\ell,t}$ Binary outage indicator for ℓth transmission line. $a = 1$ means outage; $a = 0$ otherwise.

$p^G_{g,h}(\omega)$ Scheduled active power of gth conventional generator in hth hour under scenario ω [MW].

$p^R_{r,h}(\omega)$ Scheduled active power of rth renewable generator in hth hour under scenario ω [MW].

$r_{d,h}(\omega)$ Scheduled load shedding in dth demand in hth hour under scenario ω [MW].

$f_{\ell,h}(\omega)$ Active power flow through ℓth transmission line in hth hour under scenario ω [MW].

$\delta_{n,h}(\omega)$ Phase angle at nth bus in hth hour under scenario ω [rad].

I. Introduction

The investigation on the operational patterns and economics of a power system with high renewable penetration attracts tremendous attention from both academia and industry. Despite the clean energy and zero operational cost that renewables have, the stochastic generation pattern of renewables challenges the reliability of power systems, especially when the grid requires a high level of reserves for contingencies. Thus, albeit the investment of renewables is greatly encouraged by the U.S. government and broader community [1], the system planning under high-level penetration of renewables needs to be further investigated.

Heretofore, researchers have done extensive work on the uncertainty-based system planning [2]–[5]. On the one hand, generators and lines are the biggest reliability concern from an ISO's perspective, which has been omnipresent in many operation problems, *e.g.*, unit commitment problem [6] and transmission planning problem [2]. This type of uncertainty can be effectively tackled by robust optimization (RO), as described in [2] and [7]. The basic idea of RO is to find the worst-case scenario and then make preventative decisions, which in turn also makes the solution highly conservative. On the other hand, the uncertain nature of renewable energy and elastic demand is another challenge for the system planner. A significant amount of literature (*e.g.*, [8] and [9]) also adopts RO to deal with this type of uncertainty by setting conservative boundaries. However, comparatively, stochastic programming (SP) yields less conservative solutions than RO, and we can leverage the historical data to generate scenarios. Particularly, for a large system with high renewable penetration, the spatiotemporal correlations of renewables and demand can be accurately captured by scenario generation, *e.g.*, Monte-Carlo simulation together with a multi-stage scenario tree [10].

Since the uncertainties are multiplex in today's transmission networks, *e.g.*, the binary status of generator/line and continuous generation/load volatility, the combination of SP and RO is a promising formulation with higher reliability. Recently, many works (*e.g.*, [3], [5] and[11]) also consider the hybrid SP and RO formulation with only the continuous formulation, which provides a limited evaluation on uncertainties. There are few works considering the contingencies of power system components and stochastic generation together in a generation and transmission expansion problem (G&TEP), especially under high renewable penetration levels.

The co-optimization of G&TEP under multiplex uncertainties is a nontrivial problem due to its non-convexity and uncertainties involved. A. Moreira *et al.* [4] integrated binary contingencies with continuous uncertainties but did not explore the improvement of algorithms to facilitate the solution. L. Gallego *et al.* [12] used heuristic algorithms to avoid solving an intricate model but obtained local optima. It is not trivial, however, to evaluate how far this obtained optimum is away from the global optimum. We argue that for a large-scale economic assessment like the power system planning, which may involve multiple planning years and hence does not need to be solved in real time, it would be better if we can have a guaranteed global optimum. This can be viewed as an advantage of our proposed method compared with those intelligent methods. Besides, most of the existing studies solve single-year planning problems [11], [13], [14], which do not capture certain aspects of power system planning such as the timing of the commissioning or retirement of a generator. In our study, however, the objective is to investigate whether an ultra-high renewable installation in the system is beneficial or not. Since the renewable units are widely considered as units without variable operation costs [15], the economic payback of renewable investment may outperform the conventional units over the long term. Hence, to ensure a fair comparison between different generation technologies, we also take the annuitized investment cost into the consideration and intend to uncover the cost-effectiveness of renewable investment, especially under an ultra-high renewable penetration case.

In order to investigate a 100% renewable system, the correlations between renewables, such as wind power in different regions, wind and solar, wind and demand, *etc.*, are also critical in system operation and planning. At any instant, the system should have enough generation from renewables to cover the demand. However, the system planner has to take a system-wide approach considering mutual support between regions and avoid generation shortage at a particular location. Besides, adequate transmission capacity, even in a degraded N-k line outage scenario, is paramount for inter-regional energy exchange to maintain the system energy balance. The final optimal G&TEP investment ought to consider an optimal generation mix and long-term renewable generation futures. Some researchers (*e.g.*, [14]) argue that 100% is less profitable than a generation mix with conventional units, but this is actually sensitive to the system settings.

In this paper, we propose an effective modeling and solution approach for G&TEP towards 100% penetration of renewables. The column-and-constraint generation (C&CG) method developed by [16] can decompose our problem to a master problem and a subproblem. This method has been adopted in numerous studies like [3], [6], [11], and become a mainstream solution methodology for multi-stage RO problems. Besides, as reported in [11] and [4], solving subproblems consumes the majority of the computational time. We propose to apply the L-shaped method [17] to further facilitate the solution of subproblems, which also enables parallel computing. In the meantime, we

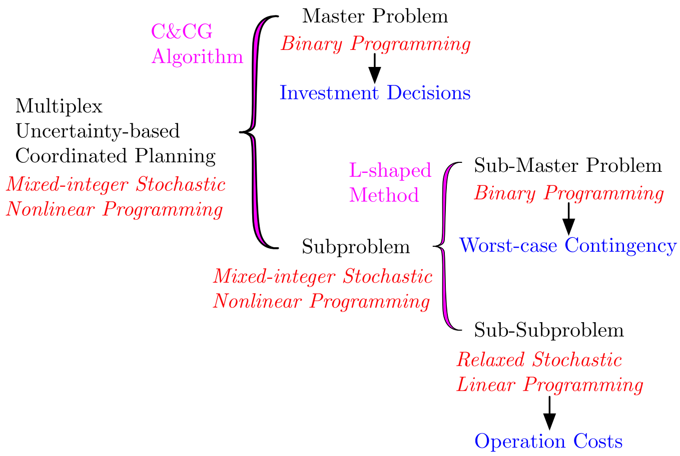


Fig. 1. Framework of the proposed methodology.

specifically consider the short-term and long-term spatiotemporal correlations between renewables and load by leveraging the Monte-Carlo sampling on regional renewables. Fig. 1, provides a general framework of our proposed methodology.

Compared with the state-of-the-art research, The main contributions of this paper are three-fold.

1) We analyze the coordinated system planning under an ultra-high level of renewable penetration and investigate the potential issues of a 100% renewable penetration on an IEEE test system and the WECC system with specific parameter settings.
2) To tackle multiplex uncertainties in the G&TEP, we propose a three-stage multi-year robust co-optimization model with a stochastic recourse. The robust counterpart captures the N-k contingency, whereas sampling-based scenarios considering spatiotemporal correlations between renewables and load realize the demand/renewable uncertainty.
3) The C&CG algorithm is leveraged to decompose the overall problem into a master-slave structure, and the stochastic subproblem is further reformulated based on the duality theory and decomposed by the L-shaped method. The computational efficiency is greatly improved, and multi-scale test cases verify the efficacy of the proposed methodology.

The rest of the paper is organized as follows: Section III describes the mathematical formulation of the G&TEP model; Section IV presents the proposed solution strategy and the detailed flowchart; case studies and a scalability test are analyzed in Section V and Section VI, respectively; Section VII summarizes this paper with several remarks.

II. Mathematical Formulation of the Generation and Transmission Planning

We provide the detailed mathematical formulation of the multiplex uncertainty-based G&TEP in this section. First, we show the deterministic equivalent form (DEF) of the three-stage formulation of the hybrid SP and RO model in (1). The term "stage" we used here refers to the mathematical structure of the proposed optimization framework.

$$\min_{\substack{x^G_{g,t},x^R_{g,t},\\ x^L_{\ell,t}}} \sum_t^T \left(\sum_g^{X^G} IC^G_{g,t} x^G_{g,t} + \sum_g^{X^R} IC^R_{r,t} x^R_{r,t} + \sum_\ell^{X^L} IC^L_{\ell,t} x^L_{\ell,t} \right)$$
$$+ \max_{a^G_{g,t},a^L_{\ell,t}} \mathbb{E}_\omega \left\{ \min_{p^G_{g,h},r_{d,h}} \sum_h^H \left[\sum_g^G OC_{g,h} p^G_{g,h}(\omega) \right.\right.$$
$$\left.\left. + \sum_d^D PC_{d,h} r_{d,h}(\omega) \right] \right\} \tag{1}$$

subject to

$$\sum_t^T \left\{ \sum_g^{X^G} IC^G_{g,t} x^G_{g,t} + \sum_g^{X^R} IC^R_{r,t} x^R_{r,t} \right\} \le B^G, \tag{1a}$$

$$\sum_t^T \left\{ \sum_\ell^{X^L} IC^L_{\ell,t} x^L_{\ell,t} \right\} \le B^L, \tag{1b}$$

$\forall * \in \{g, r, \ell\},\ \forall (*, t) \in N(*, t):$

$$\sum_{t' \neq t}^T x_{*,t'} = 0, \tag{1c}$$

$$\sum_{t' \geq t}^T y_{*,t'} = (T - t + 1) \cdot x_{*,t}, \tag{1d}$$

$$\sum_{t' < t} y_{*,t'} = 0, \tag{1e}$$

$$\sum_g^G a^G_{g,t} + \sum_r^R a^R_{r,t} + \sum_\ell^L a^L_{\ell,t} \le K_t,\ \ \forall t, \tag{1f}$$

$\forall \omega:$

$$\sum_{g|C(g)=n}^G p^G_{g,h}(\omega) + \sum_{r|C(r)=n}^R p^R_{r,h}(\omega) - \sum_{\ell|S(\ell)=n}^L f_{\ell,h}(\omega) +$$
$$\sum_{\ell|V(\ell)=n}^L f_{\ell,h}(\omega) = \sum_{d|C(d)=n}^D \{\ P_{d,h}(\omega) - r_{d,h}(\omega)\ \}, \forall n, \forall h, \tag{1g}$$

$$f_{\ell,h}(\omega) = y^L_{\ell,t}(1 - a^L_{\ell,t}) A_\ell^{-1}$$
$$\cdot [\delta_{n|S(\ell)=n,h}(\omega) - \delta_{n|V(\ell)=n,h}(\omega)],$$
$$\forall \ell \in L, \forall (t,h) \in M, \tag{1h}$$

$$-y^L_{\ell,t}(1 - a^L_{\ell,t}) FL_\ell \le f_{\ell,h}(\omega), \forall \ell \in X^L, \forall (t,h) \in M, \tag{1i}$$

$$f_{\ell,h}(\omega) \le y^L_{\ell,t}(1 - a^L_{\ell,t}) FL_\ell, \forall \ell \in X^L, \forall (t,h) \in M, \tag{1j}$$

$$-(1 - a^L_{\ell,t}) FL_\ell \le f_{\ell,h}(\omega), \forall \ell \in L \backslash X^L, \forall (t,h) \in M, \tag{1k}$$

$$f_{\ell,h}(\omega) \le (1 - a^L_{\ell,t}) FL_\ell, \forall \ell \in L \backslash X^L, \forall (t,h) \in M, \tag{1l}$$

$$p^G_{g,h}(\omega) \le y^G_{g,t}(1 - a^G_{g,t})\overline{PL}^G_g, \forall g \in X^G, \forall (t,h) \in M, \quad (1m)$$

$$p^G_{g,h}(\omega) \le (1 - a^G_{g,t})\overline{PL}^G_g, \forall g \in G \backslash X^G, \forall (t,h) \in M, \quad (1n)$$

$$p^R_{r,h}(\omega) \le y^R_{r,t}(1 - a^R_{r,t})\overline{PL}^R_{r,h}(\omega), \forall r \in X^R, \forall (t,h) \in M, \quad (1o)$$

$$p^R_{r,h}(\omega) \le (1 - a^R_{r,t})\overline{PL}^R_{r,h}(\omega), \forall r \in R \backslash X^R, \forall (t,h) \in M, \quad (1p)$$

$$\underline{\Delta}_n \le \delta_{n,h}(\omega) \le \overline{\Delta}_n, \forall n, \forall h, \quad (1q)$$

The objective function (1) formulates the investment cost of the proposed generator and transmission line expansions, plus the operating cost of scheduled conventional generators and possible load shedding.

Constraints (1a)-(1b) model the investment budget of both generators and transmission lines, which construct the first-stage feasible region. Constraints (1c)-(1e) represent that once an investment is made in year t, the component will be available for the rest of the planning horizon. Control sets $N(*, t)$ represent the mapping between the candidate units and planning time, in which we categorize the candidate units based on the investable years. These constraints formulate the constraint space for the first stage. Constraint (1f) shows the N-k criteria for both generators and transmission lines, where K_t can be adjusted to perform different contingency analyses. This constraint is in the second-stage (the second-stage objective function can be regarded as $\{\mathbf{0}^\top \mathbf{a}\}$, in which $\mathbf{a}$ is the vector of the outage indicators) formulating the uncertainty set [18]. This uncertainty set is discretely polyhedral.

For the third-stage, we first explain the relationship between t and h. Since we consider a long-term planning problem, traversing 8760 hours for a whole year renders heavy intractability in this model. Hence, we consider a 24-hour operation for the third stage as an analysis of hourly dispatch in one typical day. This operation is also adopted in [14]. Then the time index mapping set M can be described as

$$M = \{(t_1, h_1), ..., (t_1, h_{24}), (t_2, h_{25}), ...(t_2, h_{48}), ...\}.$$

We also note that this design further generalizes the application of selecting representative hours/days. When a system planner intends to select more representative hours to perform the G&TEP study, it is trivial to adjust the set M by increasing the number of hours. This also shows that our proposed framework is highly flexible and general to be expanded to the extent the system planner would like to use in G&TEP problems.

Notably, the outage indicator in our formulation is a day-based variable. As we indicated, the third-stage problem can be regarded as a 24-hour economic dispatch problem. For such operation problems, the contingency is often considered throughout the operation horizon, *i.e.*, 24 hours [19]. Practically, 94% of the planned and operational outage have a duration of over 2 hours and 34.6% of them are over 48 hours [20]. In some works of system planning, the horizon-long contingency is also adopted [5]. For the sake of simplicity, since we choose one representative day for one year with multiple uncertain scenarios, we slightly abuse our notation and use a year-based index to represent a day-based variable, *e.g.*, $a^G_{g,t}$. Besides, for the third-stage operation problem, we choose using one representative day with multiple scenarios to investigate the short-term correlation between regional renewables and load, which can be better captured by hourly operations, as also adopted in [14].

Particularly, constraint (1g) is the nodal balance constraint, and constraint (1h) defines the DC line flow equations. Note that here, we slightly abuse the notation that the per unit reactance A_ℓ should be normalized under the system base MVA. Constraints (1i)-(1l) are line flow limitation constraints for both candidate and existing lines. Constraints (1m)-(1p) are the generation capacity constraints for both candidate and existing conventional and renewable generators. Note that the renewable generators can be dispatched, and thus we permit the renewable curtailment. Constraint (1q) shows the phase angle limitation. We can find that all of the third-stage variables are associated with the scenario index ω for different realizations of renewable generation and load. Note that the binary variables and the variable multiplications such as in (1h), (1i), (1j), (1m) and (1o) render this model mixed-integer nonlinear.

The proposed model (1) is a very general framework for the hybrid stochastic and robust optimization-based system planning and can be easily adjusted or expanded to include more constraints regarding various research directions or industrial applications. In our case studies, for example, we do not have any existing generator and we can hence disregard constraints (1n) and (1p) in the formulation. We also modify (1) to incorporate unique case settings such as the regional N-1 contingency and multiple investments in one candidate bus, which can be achieved by replacing constraint (1f) with (1r) and adding constraint (1s) respectively as shown below.

$$\sum_g^{G^z} a^G_{g,t} + \sum_r^{R^z} a^R_{r,t} + \sum_\ell^{L^z} a^L_{\ell,t} \le K_t, \quad \forall t, \forall z \in Z, \quad (1r)$$

$$\sum_{g|C(g)=n^{cg}}^{G} x^G_{g,t} + \sum_{r|C(r)=n^{cr}}^{R} x^R_{r,t} \le 2, \quad \forall t, \forall n, \quad (1s)$$

where z denotes the region indicator, Z is the total region set, n^{cg} denotes the candidate bus for thermal generators, and n^{cr} denotes the candidate bus for renewable generators.

III. Problem Decomposition and Solution Strategy

The model described in Section III presents an intractable three-stage mixed-integer nonlinear problem with stochastic recourse. However, we can decompose the original problem into a structure of mixed-integer linear programming (MILP) master problem and mixed-integer bilinear subproblem, which can be solved iteratively based on the theory of the C&CG algorithm [16]. Linear relaxation techniques can tackle the nonlinearity of the subproblem. Besides, the subproblem can be further decomposed by the L-shaped method [17].

Fig. 2, depicts the overall solution workflow. Generally, we adopt the C&CG algorithm to decompose the problem into a master-slave structure, and the L-shaped algorithm then further decomposes the subproblem by different scenarios. We will

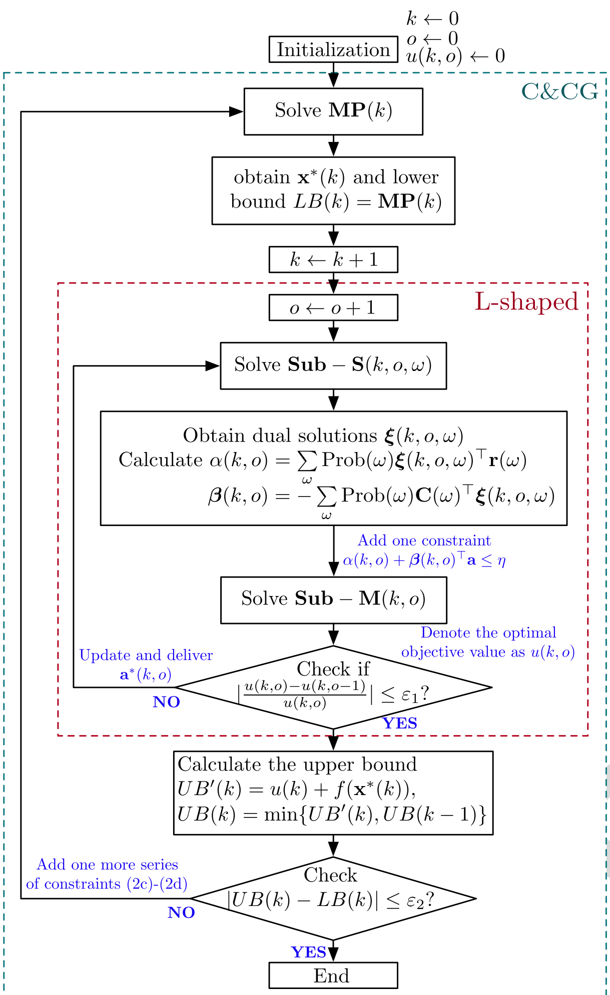


Fig. 2. Workflow of the proposed algorithm.

explain the procedure and notations of Fig. 2, in the problem formulation. To be more concise, we use compact form below.

A. Master Problem

We formulate the master problem in the C&CG procedure as in (2). $\mathbf{x}$ denotes the first-stage variables, $\mathbf{z}$ denotes the third-stage variables with iteration index k', which encloses all the past iterations before the current iteration k. $\mathbf{IC}$ and $\mathbf{PC}$ represent the vectors of investment costs and penalty costs, whereas $\mathbf{p^G}$ and $\mathbf{r}$ denote the vectors of generator dispatch and load shedding. ϕ is an auxiliary variable that formulates a relaxed lower bound of the second-stage problem. The constraints for third-stage variables are (2d). Note that the second-stage variables $\mathbf{a}$ are fixed here, which are delivered from the subproblem. Hence, the master problem now becomes a deterministic MILP problem that can be efficiently tackled via off-the-shelf solvers. Here, $\mathbf{x}$ includes the first-stage variables, *i.e.*, $x^G_{g,t}, x^R_{r,t}, x^L_{\ell,t}, y^G_{g,t}, y^R_{r,t}, y^L_{\ell,t}$ and $\mathbf{z}$ includes the third-stage variables, *i.e.*, $p^G_{g,h}, p^R_{r,h}, r_{d,h}, f_{\ell,h}, \delta_{n,h}$.

$$\mathbf{MP} = \min_{\mathbf{x},\mathbf{z}} \mathbf{f}(\mathbf{x}) + \phi \qquad (2)$$

subject to

$$\text{Constraints (1a)-(1e)} \qquad (2a)$$

$$f(\mathbf{x}) = \mathbf{IC}^\top \mathbf{x}, \qquad (2b)$$

$$\phi \geq \mathbb{E}_\omega \left\{ \mathbf{OC}^\top \mathbf{p^G} + \mathbf{PC}^\top \mathbf{r} \right\}, \ \forall k' \leq k, \qquad ((2c))$$

$$\text{Constraints (1g)-(1q)}, \ \ \forall k' \leq k \qquad (2d)$$

B. Subproblem

The subproblem of the C&CG procedure is constructed from the second- and third-stage problems, *i.e.*, the problem determining the worst-case scenario of contingency. Here, $\mathbf{a}$ includes the second-stage variables, *i.e.*, $a^G_{g,t}, a^R_{r,t}, a^L_{\ell,t}$.

$$\max_{\mathbf{a}} \ \mathbb{E}_\omega \left\{ \min_{\mathbf{p},\mathbf{r}} \ \mathbf{OC}^\top \mathbf{p^G} + \mathbf{PC}^\top \mathbf{r} \right\} \qquad (3)$$

subject to

$$\text{Constraints (1f)-(1q).} \qquad (3a)$$

Note that the first-stage variables in the constraints become fixed parameters obtained from the master problem in the previous iteration. Since the inner minimization problem of (3) has linear programming characteristics, according to the strong duality theory, it is equivalent to rewrite (3) to its dual form (4), after giving an objective handle for the second-stage.

$$\mathbf{Sub} = \max_{\mathbf{a}} \ \ \mathbf{0}^\top \mathbf{a} + \mathbb{E}_\omega \left\{ \max_{\boldsymbol{\pi}} \ \mathbf{Q}(\mathbf{a}, \boldsymbol{\pi}) \right\} \qquad (4)$$

subject to

$$Q(\mathbf{a}, \boldsymbol{\pi}) \in \boldsymbol{\Gamma}_{\mathbf{a},\boldsymbol{\pi}}, \qquad (4a)$$

where $Q(\mathbf{a}, \boldsymbol{\pi})$ represents the dualized objective function, $\boldsymbol{\pi}$ is the vector of all dual variables in (3) and $\Gamma_{\mathbf{a},\boldsymbol{\pi}}$ consists of the constraint space. For better demonstration, problem (4) is further altered to a minimization problem (5) where we rewrite $\Gamma_{\mathbf{a},\boldsymbol{\pi}}$ in the form of (5a)-(5c). Specifically, (5a) constructs the uncertainty set, (5b) formulates the operational constraints, and (5c) ensures the dual feasibility of the KKT condition.

$$\mathbf{Sub} = \min_{\mathbf{a}} \ \ -\mathbf{0}^\top \mathbf{a} \ + \ \mathbb{E}_\omega \left\{ \min_{\boldsymbol{\pi}} \ -\mathbf{Q}(\mathbf{a}, \boldsymbol{\pi}) \right\} \qquad (5)$$

subject to

$$\text{Constraint (1f)}, \qquad (5a)$$

$$\mathbf{D}(\omega)\boldsymbol{\pi} \leq \mathbf{r}(\omega) - \mathbf{C}(\omega)\mathbf{a}, \qquad (5b)$$

$$\boldsymbol{\pi} \geq 0. \qquad (5c)$$

Note that in the objective function, $Q(\mathbf{a}, \boldsymbol{\pi})$ includes the first-stage decisions $\mathbf{x}^*$ determined from the master problem. This subproblem shows a typical two-stage stochastic mixed-integer bilinear minimization structure. Furthermore, linear relaxation techniques, *e.g.*, the Big M method (see Appendix A), can effectively relax the bilinear parts without sacrificing accuracy. As the binary nature of the integer variables tightens the convex relaxation, we can ensure the accuracy of the obtained solution. According to the stochastic L-shaped method [17], the two-stage mixed-integer stochastic program with linear recourse and finite support is also decomposable, as shown in the following.

1) L-shaped Master Problem:

$$\mathbf{Sub-M} = \min_{\mathbf{a}} \quad -\mathbf{0}^{\top}\mathbf{a} + \eta \tag{6}$$

subject to

$$\text{Constraint (1f)}, \tag{6a}$$

$$\alpha_{o'} + \boldsymbol{\beta}_{o'}^{\top}\mathbf{a} \le \eta, \quad \forall \mathbf{o}' \le \mathbf{o} \tag{6b}$$

In the L-shaped master problem, η is an auxiliary variable, $\alpha_{o'}$ and $\boldsymbol{\beta}_{o'}$ are subgradients computed from the dual of the L-shaped subproblem, which will be discussed in the next subsection. Note that this formulation is for the single-cut L-shaped algorithm.

2) L-shaped Subproblem:

$$\mathbf{Sub-S}(\omega) = \min_{\boldsymbol{\pi}} \quad -\mathbf{Q}(\mathbf{a}^*, \boldsymbol{\pi}) \tag{7}$$

subject to

$$\mathbf{D}(\omega)\boldsymbol{\pi} \le \mathbf{r}(\omega) - \mathbf{C}(\omega)\mathbf{a}^* : \qquad \boldsymbol{\xi}(\omega), \tag{7a}$$

$$\boldsymbol{\pi} \ge 0. \tag{7b}$$

When the L-shaped master problem yields an optimal solution of $\mathbf{a}^*$, the subproblem receives this solution and solves the operation problem. Let $\boldsymbol{\xi}^*(\omega)$ denote the optimal dual solution of the operation constraints (7a). After solving ω individual L-shaped subproblems, which can be done in parallel, the subgradients for each iteration can be computed as follows.

$$\alpha_o = \sum_{\omega} \text{Prob}(\omega)\boldsymbol{\xi}_o^*(\omega)^{\top}\mathbf{r}(\omega),$$

$$\boldsymbol{\beta}_o = -\sum_{\omega} \text{Prob}(\omega)\mathbf{C}(\omega)^{\top}\boldsymbol{\xi}_{\mathbf{o}}^*(\omega).$$

Afterwards, the L-shaped master problem receives an optimality cut (6b). Finally, the inner iteration loop for the L-shaped procedure determines the final worst-case contingency $\mathbf{a}^*$ and sends it back to the C&CG master problem. As we enable load shedding, the feasibility of all problems is guaranteed, and the feasibility cut is therefore negligible.

The proposed algorithm guarantees its convergence by the finite extreme points of the uncertainty set and finite support in the second-stage stochastic recourse according to the convergence analysis [16] and [17]. For the convergence speed, according to Fig. 2, the C&CG procedure does not influence the convergence of the embedded L-shaped method, which means we still retain the fast convergence of the C&CG method [16]. The convergence of the C&CG part relies on the convergence of the L-shaped algorithm, which is guaranteed by the finite support in the stochastic recourse. The motivation of leveraging the L-shaped method for the subproblem is that solving the original subproblem consumes the majority of the computational time, as reported in [4] and [11]. We enjoy the merit of parallel computing to facilitate solving subproblems when the L-shaped method is applied.

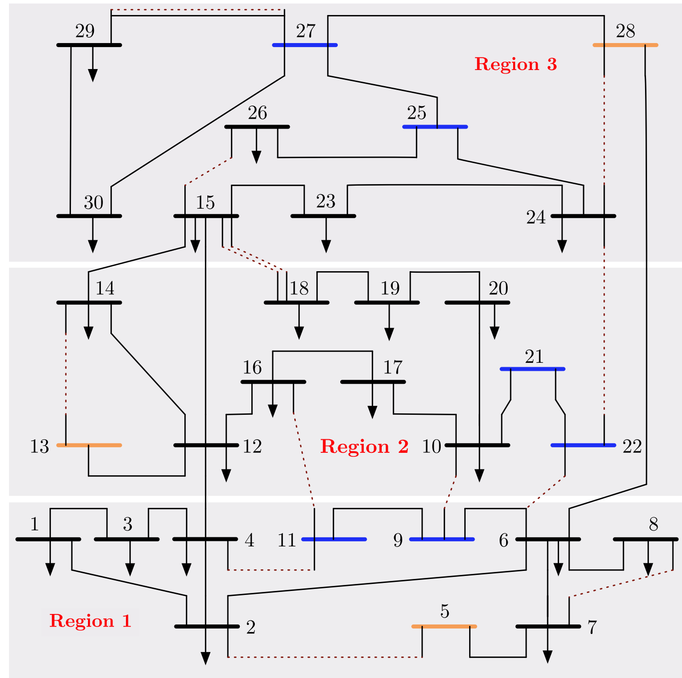


Fig. 3. System topology of the modified IEEE 30-bus system.

TABLE I
DATA FOR CANDIDATE GENERATION UNITS

Generation Type	Operating cost [\$/MWh]	Maximum capacity [MW]	Overnight capital cost [M\$]
Wind	0	300	390
Solar	0	150	170
Thermal	41.2	300	67.8
Trans. Line	0	600	50

IV. CASE STUDIES

This section provides results and discussions for case studies on a modified IEEE 30-bus system based on [21]. We aim to investigate the benefit of ultra-high renewable penetration towards 100% in a long-term planning problem. We implement all of the experiments in GAMS 25.0.3 [22] with CPLEX 12.8 and run it on a 2.60GHz Windows PC with a 6-core Intel i7 CPU and 8GB RAM. We also leverage a GAMS-embedded parallel computing tool, *i.e.*, Gather-Update-Solve-Scatter (GUSS) [23], for the L-shaped subproblem to improve the computational efficiency.

To illustrate our proposed methodology, we modify the IEEE 30-bus system to have reduced transmission lines and no existing generators. Fig. 3, provides the detailed topology, where the blue and orange buses indicate the candidate locations for building wind and solar generators, respectively. Thermal generators can be invested in any load bus. The dashed lines indicate the candidate transmission lines. Table I shows the investment information, whose reference can be found in [24]–[26]. The modified IEEE 30-bus system has a total daily peak power

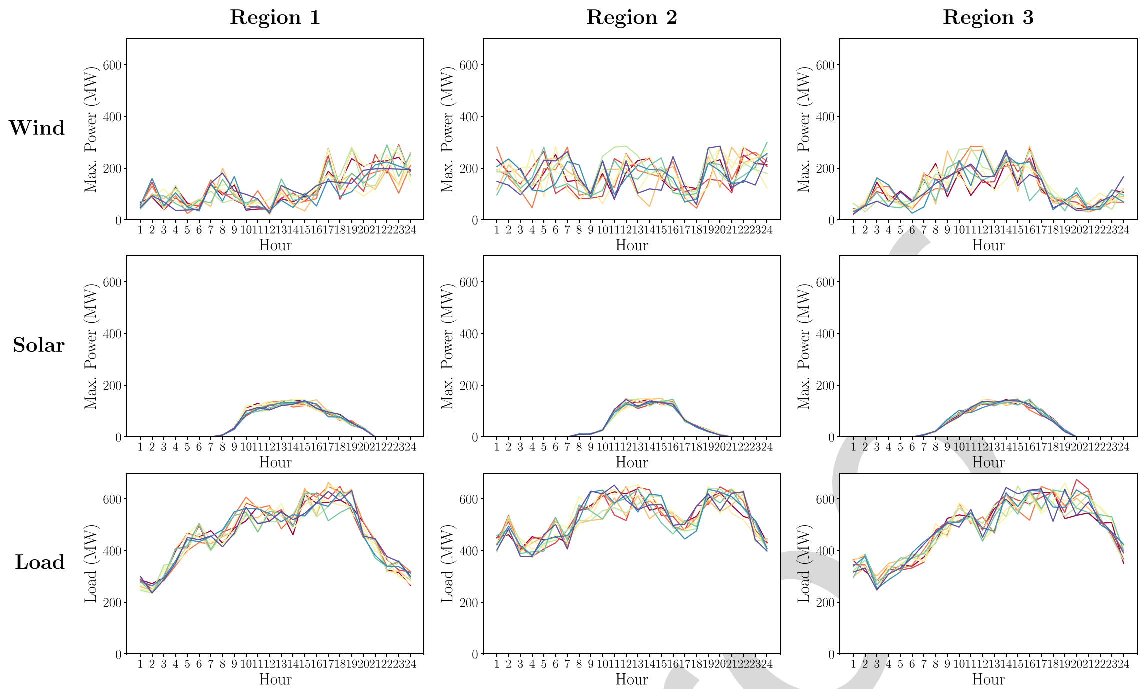


Fig. 4. Regional scenario (10 for each).

demand as 2,000 MW, distributed to the three regions. And we set the load shedding penalty cost as $1,000/MWh.

A. Scenario Generation & Reduction

There are three regions in the system, where different scenario sets of renewable and demand are applied. To verify the investment performance considering the spatiotemporal correlation between renewables and load, we create the scenario set for each region individually by using Monte-Carlo simulation on three different datasets of renewable output from [27]. Fig. 4, depicts the regional 24-hour sequential wind, solar, and demand scenarios. The average capacity factors of wind and solar are 45.14% and 28.86%, respectively. We also assume that the two neighbored wind buses in Fig. 3, share the same wind output time series, and one can easily simulate scenarios with respect to different time series for more detailed spatiotemporal studies.

For the spatiotemporal correlations between renewables and load, after sampling on three different datasets with respect to three different regions, we apply the Fast Backward/Forward method embedded in the GAMS SCENRED toolbox to reduce the scenario number according to the balance between the number of scenarios and the solution accuracy. This strategy is generally an approximation of the complete scenario tree benchmarked by different criteria such as the Fortet-Mourier metric [28] and L_r-distance [29], which have been widely adopted in many stochastic system planning works *e.g.* [30] and [31]. To capture the spatiotemporal correlations between renewables and demand, we use three temporal datasets for real renewable outputs and demand in three forecast zones in the ERCOT area as a basis for our Monte-Carlo sampling to create the scenario sets. With both the demand and renewable uncertainties being taken into consideration, the proposed G&TEP framework can yield planning results according to the correlations between multiplex uncertainties. It is also imperative to indicate that a more accurate and efficient scenario generation/reduction technique concerning the spatiotemporal correlation between uncertainties is of great future research interest, to which our proposed framework in this paper can easily adapt.

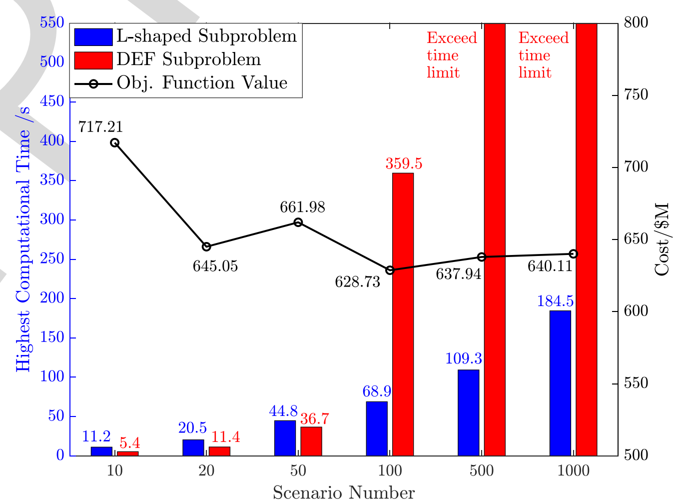


Fig. 5. Comparison for different sizes of scenarios.

To balance the tradeoff between the accuracy and tractability, we test different scenario sets by running a one-year planning problem where two respective investments on solar and wind generators are mandatory, as shown in Fig. 5, and the highest computational time stands for the highest time of solving subproblems among all iterations. Since the difference in objective function values between 100 scenarios and 1000 scenarios is within 2%, we argue that the set of 100 scenarios is accurate enough to perform the following analyses. Besides, Fig. 5, also validates the necessity of using decomposition techniques for solving stochastic subproblems, as the computational burden grows exponentially with the increasing number of scenarios in the DEF problem as in [3] and [11], but grows linearly in the L-shaped problem in our work.

B. *Case Studies: Towards 100% Renewable Penetration*

We design three cases to show the pathways to achieve 100% renewable penetration. We apply the N-1 criterion to each region. Each candidate bus can build two generators in each year. The algorithmic convergence gap ε_2 is set to be 0.1%, and the solver's MIP gap is set as 0.01%.

- *Case 1:* \$900M generation investment budget (6yrs);
- *Case 2:* \$9,000M generation investment budget (6yrs);
- *Case 3:* \$9,000M generation investment budget. Thermal units can only be installed in the first year. All the thermal units will phase out by 20% capacity per year (6yrs) with a salvage income.

To be more practical, for thermal units, we consider a 5% annual inflation rate of the fuel price. For renewable energy, the investment cost has a discount factor of 5% per year as the technology develops. The salvage price of phased out thermal generators is 40% of the investment cost. The total system demand also increases by 5% per year. Table II shows the investment decisions, costs and system information obtained from the three cases.

For the salvage income, we directly add a salvage income term in the objective function of (1). The salvage of generators includes the sales of the salvageable parts of the unit, recycling worn-out equipment, and reutilizing the designated real estate [31]. According to [32], the salvage value for generators is calculated based on a linear relationship with the proportion of the used life and the remaining life, resulting the following equation:

$$S = C_{replace} \cdot \frac{R_{remain}}{R_{component}}$$

in which $C_{replace}$ is the replacement cost that is about 80% of the initial investment cost, R_{remain} is the remaining life and $R_{component}$ is the component lifetime. We also refer to [33] for the lifetime of a pulverized coal power plant as 30 years. Thus, in year 6, the remaining lifetime of such unit is 24 years, and thus the salvage income should be $80\% \cdot (24/30) = 64\%$ of the investment cost. Considering the demolition cost and the personnel cost, we set the salvage value for a thermal unit in our study as 40% of the investment cost. We will revisit *Case 1* and *Case 2* in subsection E by considering the investment cost annuitization to further assess the economic aspects of renewable installations.

1) General Investment Plan and Operation: For the generation investment, on the one hand, since we consider the N-1 criterion in each region, the system planner invests in a large amount of generation capacity, especially when the renewable is installed (*e.g.*, the total capacity in *Case 3* is 152.9% higher than the peak load). While the system planner has already known the demand increase rate for the later years, most of the new generation capacity is invested in the first year to meet the peak demand growth in the following years. However, since the renewable installation is limited compared with the thermal installation, and considering the yearly increasing system demand and the unit contingencies every year, there are still investments after the first year especially in *Case 2* and *Case 3* to ensure the system reliability. Another noteworthy issue is renewable curtailment. The peak curtailment in *Case 3* is 18.29% when the average curtailment across the scenarios also reaches 12.11%. They are comparatively small since the excess renewable power can be used to support other regions. Future works considering using the curtailed energy to provide ancillary services and storage charging can be envisioned.

On the other hand, for the transmission investment, new transmission lines tend to be built in the first year to prepare for the worst-case contingency in the subsequent years. The new transmission build-out is mostly to accommodate the injections from the new generators. Notably, lines 6-8, 12-15, and 6-28 are reported to be the ones with the most frequent outage under the worst-case contingency. Thus, in the investment portfolio, lines 7-8, 6-22, and 22-24 are always built in the three cases to aid power transmission. The system finally achieves the 100% renewable penetration level in the 6$^{\text{th}}$ year in *Case 3*. The 100% is based on the generation capacity percentage.

TABLE II
INVESTMENT PORTFOLIO REPORT FOR THE 30-BUS SYSTEM

Gen. Investment*	*Case 1*	*Case 2*	*Case 3*
Year 1	2, 4, 6, 8, 10, 12, 14, 15, 18, 24, 29, 30	2, 4, 5, 5, 10, 9, 14, 15, 18, 21, 22, 24, 27, 29	4, 5, 5, 9, 9, 11, 11, 16, 21, 21, 22, 22, 25, 25, 27, 27, 30
Year 2	None	21	9, 13, 21, 22, 25, 25
Year 3	None	None	None
Year 4	None	None	None
Year 5	None	None	None
Year 6	None	None	None
Line Investment	*Case 1*	*Case 2*	*Case 3*
Year 1	4-11, 7-8, 6-22, 9-10, 15-26, 22-24,	4-11, 7-8, 6-22, 9-10, 15-18, 15-26, 22-24	4-11, 7-8, 6-22, 9-10, 15-18, 15-24, 15-26, 22-24
Year 2-6	None	None	None
Gen. Invest. Cost	\$813.6M	\$2,812.9M	\$7,237.4M
Gen. Opera. Cost	\$5,976.81M	\$3,542.70M	\$168.93M
Line Invest. Cost	\$300M	\$350M	\$400M
Load Shed Cost	0	0	\$1,079.43M
Salvage Income	0	0	\$81.36M
Total Cost	\$7,090.41M	\$6,705.60M	\$8,754.40M
Avg. Load Shed Portion	0%	0%	0.96%
Avg. Renewable Curtailment	0%	7.75%	12.11%
Final Renewable Percentage	0%	42.86%	100%

* The numbers indicate the buses built with thermal, wind and solar.

By comparing *Case 1* and *Case 2* in Table II, we can find that if we have a sufficient investment budget at the beginning, it is more profitable to invest in a proper generation mix of thermal, wind and solar technologies from the system perspective. Since we consider a long-term investment plan, the operation cost for 6 years in *Case 1* far outnumbers the investment cost, whereas the renewable generators only have a one-time cost for installation.

By comparing *Case 2* and *Case 3* in Table II, we find that the operation cost is further greatly decreased in *Case 3*, but the investment cost and the load shedding cost increase significantly, which results in a higher total cost. However, as the investment cost of renewable technology decreases yearly, in the long-term planning, the system will tend to have higher renewable penetration approaching 100%, which will be detailed in subsection D.

In the current setting, we do not consider the energy storage and the time-shifting demand response since we would like to investigate the cost-effectiveness of renewables only from their own economic aspects. However, it is imperative to state that considering 100% renewable energy without any energy storage or demand response is not beneficial for all circumstances. It can be envisioned that energy storage and demand response can greatly reduce the need for over-installation of renewable generation.

2) Interregional Support With Spatiotemporal Correlations: The spatiotemporal correlations between renewables and regional differences of the uncertainty also influence the investment portfolio of renewables. When we do not consider any storage device or time-shifting demand response in our model, there could be a large amount of load shedding in a particular region when the renewable generation in that region is very low. However, such a situation can be largely mitigated with adequate transmission capacity with the other regions since the variation of renewable generation output from each region tends to cancel out over a large geographical area to provide a more stable output. Hence, by using the proposed scenario generation and the optimization framework, the new generators are built in the way that they can provide interregional energy support when needed, based on their spatiotemporal correlations.

On the one hand, from the long-term investment perspective, in Fig. 4,, the diurnal wind output in Region 1 is small, which incentivizes the investment of the solar unit at bus 5 and reduces the investment of wind units in the region. Besides, in *Case 3*, the lines connecting Region 2 and Region 3, *i.e.*, line 15-18 and line 22-24, are all constructed to help Region 3 cover its demand as the wind output in Region 3 in a long run is relatively lower compared with the other regions.

On the other hand, from the short-term operation perspective, in the 6th year of *Case 3*, we particularly analyze one operation scenario. Fig. 6, depicts the active power flow in the tie-lines with Region 1, *i.e.*, lines 9-10, and 6-22. It can be seen that the power support from the other regions reaches the peak value at hour 10 due to the lack of local wind power in Region 1, but Region 1 can export active power to the other regions at night when the wind output increases. This phenomenon validates the interregional support when the spatiotemporal correlations between renewables and load are present.

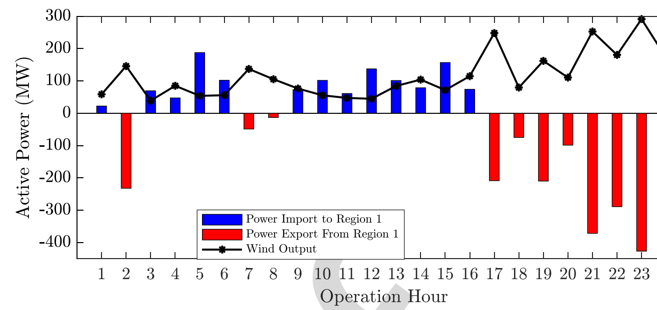


Fig. 6. Power transfer from/to Region 1: one scenario.

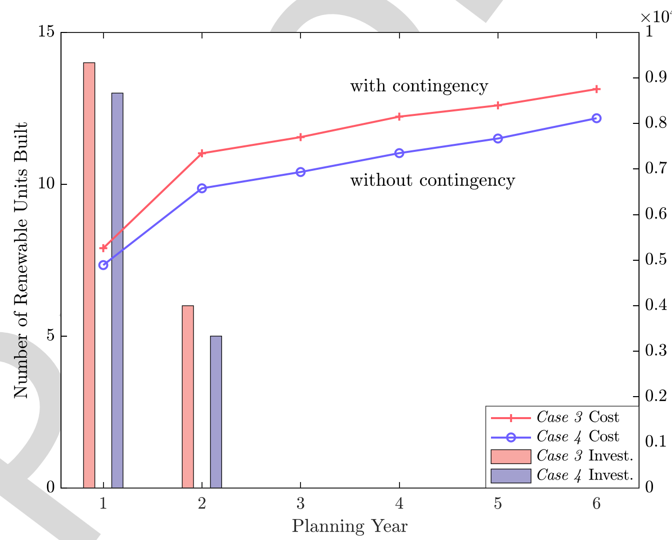


Fig. 7. Comparison in the contingency criterion.

C. Case Studies: Contingency Criterion

Next, we elaborate on the contingency criterion and the impact of the worst-case contingency by comparing the following *Case 4* with the previous *Case 3*.

- *Case 4*. We apply no contingency criterion in all regions, *i.e.*, the robust counterpart in the model is omitted. The other settings are the same as *Case 3*.

Fig. 7, demonstrates the generator's installation details of these two cases. It can be seen from the results that when the N-1 contingency criterion is applied for each region in *Case 3*, the system needs to have more renewable generators installed to secure the demand. Particularly for the worst-case contingency, the system is prone to increase the investment in generators to provide more contingency reserves, which is also one of the reasons of the huge installations in *Case 3*.

D. Case Studies: Long-Term Cost-Effectiveness

To further elaborate the cost-effectiveness of long-term investment in renewables, especially for the 100% renewable penetration, we carry out two case studies of a 12-year expansion in the following:

- *Case 5:* $12,000M generation investment budget (12yrs);

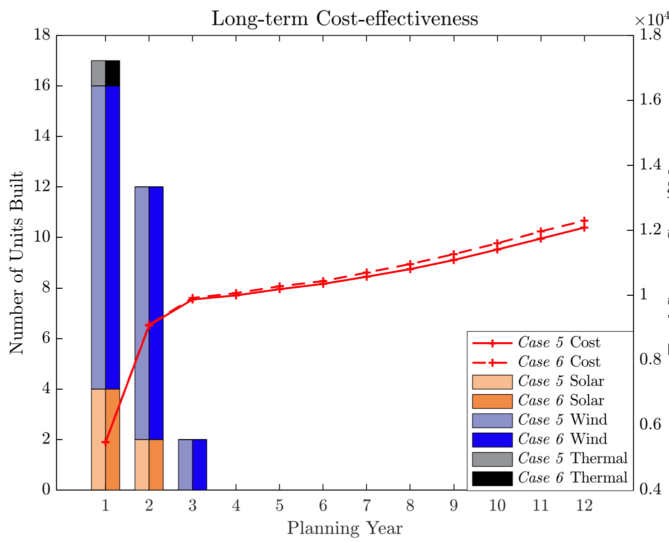


Fig. 8. Investment plans and system cost of the long-term planning.

- *Case 6:* \$12,000M generation investment budget. Thermal units can only be installed in the first year. All the thermal units phase 10% capacity out per year (12yrs).

It is straightforward that the 100% renewable penetration is accomplished in the 11$^{\text{th}}$ year in *Case 6*. We enable the transmission expansion between any two connected buses for the annually increasing demand, and the other system settings are the same as in the previous case studies. Fig. 8, depicts the investment plans and system costs. Accordingly, the salvage value of conventional generators is set as 25% of the investment cost in this case.

In *Case 5*, the installation of one thermal unit is mainly to balance the load shedding caused by stochastic generation, even when the interregional support from renewables has already largely mitigated this issue. And in *Case 6* when the condition changes that the thermal units are no longer able to fully provide energy, the system planner is still prone to invest one thermal unit due to the high load shedding cost. Though the 100% renewable case, *i.e.*, *Case 6*, still has a slightly higher cost than *Case 5*, we argue that the final cost difference highly depends on how we choose the system parameters. Nonetheless, from these two cases, we can see the potential of renewable generation's long-term cost-effectiveness. In *Case 5* when the phasing-out is not allowed, the system in the 12$^{\text{th}}$ year still reaches 96.43% renewable penetration, which implies the long-term economic payoff is higher than conventional units. It should also be noted that different settings of renewable-based G&TEP studies could lead to different findings (*e.g.*, [14]) but the proposed general framework still applies.

E. Case Studies: Investment Annuitization

In many investment studies, investors often consider annuitization of the investment cost to distribute the investment over the planning horizon. Hence, in this section, we illustrate the effect of considering the annuitization of the investment cost with the same settings of *Case 1* and *Case 2*. We keep the same investment budget as in *Case 1* and *Case 2*, and consider a general annual interest rate i of the investment as 6.04% computed from the weighted average cost of capital (WACC) of 5.7% [34]. We adopt the investment annuitizing method from Appendix A in [34] as we first calculate the annual discount factor (DF_t) by

$$DF_t = \frac{1}{(1+i)^t},$$

then the annualized capital cost (ACC_g) and the discounted investment cost ($IC_{g,t}$) used in the objective function can be computed as

$$ACC_g = OCC_g \cdot \frac{i \cdot (1+i)^{LT_g}}{(1+i)^n - 1},$$

$$IC_{g,t} = ACC_g \cdot \sum_{t' \leq \min\{LT_g, T^{\text{remain}}\}} DF_{t'},$$

where OCC_g denotes the overnight capital cost, LT_g denotes the generator lifetime, and T^{remain} denotes the remaining time of the planning horizon. The annual increase in thermal units' fuel cost and demand still applies.

TABLE III
INVESTMENT PORTFOLIO WITH THE INVESTMENT ANNUITIZATION

Investment*	*Revisited Case 1*	*Revisited Case 2*
Year 1	10 0 0 6	5 6 1 6
Year 2	0 0 0 0	0 3 0 1
Year 3	1 0 0 0	0 0 1 0
Year 4	1 0 0 0	1 0 0 0
Year 5	0 0 0 0	0 1 0 0
Year 6	0 0 0 0	0 0 0 0
6-yr Invest. Payment	\$762.67M	\$4,153.67M
Line Invest. Cost	\$300M	\$350M
Gen. Opera. Cost	\$5,978.10M	\$2,163.77M
Total Cost	\$7,040.78M	\$6,667.44M

* The integrals denote the numbers of invested thermal, wind, solar generators, and line.

Table III tabulates the results of the revisited *Case 1* and *Case 2* with considering the annuitized investment cost. In both cases, there is no load shedding. Thanks to the annuitization, the investment of generators can now be distributed to later planning years. Compared with the previous *Case 1* and *Case 2*, we find that both total costs do not differ much, and the results in *Case 1* remain nearly the same except for some investments distributed to later years. But in *Case 2*, the renewable installation is more prioritized, as the investment annuitization makes the renewable investment more competitive with thermal investments. However, we can still draw similar conclusions that a proper generation mix of conventional and renewable technologies renders a more cost-effective planning portfolio if the budget allows for renewable installations. As the annuitization method has already taken the generators' lifetime into account [34], we do not apply the technique of considering salvage incomes [31] in the revisited cases. Leveraging the annuitization method can make the investment in multi-year planning more comparable,

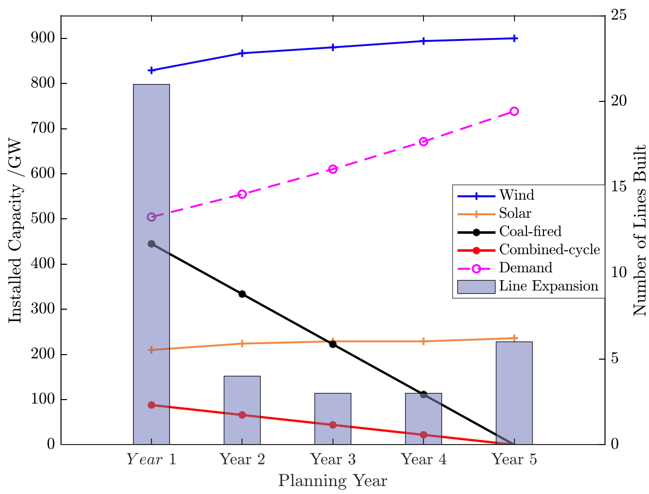


Fig. 9. Installed capacity for the WECC system.

which should be included in a more realistic G&TEP study. Also, it is noteworthy that the performance is highly sensitive to the selected WACC determined by the tax rate and expected returns on equity and debt.

V. Scalability Test

To validate the scalability of the proposed framework, we modify the WECC 243-bus system based on [35] for G&TEP studies and carry out the scalability test. Four generation technologies are included, *i.e.*, coal, combined-cycle, wind, and solar. We carry out a regionally N-1 contingency-constrained 5-year planning problem with the hourly operation. The number of combinatorial scenarios, in this case, is further reduced to 10 to reduce the computational burden. And we enable the phasing-out of conventional units as in *Case 3*.

Fig. 9, depicts the scheduled installed capacity for each generation technology and transmission line, where we can observe that the investment in solar and wind generators grows sharply in the first year and keeps increasing, due to the 25% annual phasing-out rate of the conventional generators. The computational time of the WECC simulation is 27 hours, which is comparably reasonable concerning the scale of the multi-year stochastic and robust planning problem.

VI. Discussions

To clarify the scope of our work and how it should be used as a reference for both academia and industry, we provide several discussions on this paper.

1) Though the day-based contingencies have already been adopted in this paper and other works, *e.g.*, [5], [19], the contingencies of power system components can be evaluated in a practically smaller time resolution.
2) Using only a few representative days has been found to be not enough, and more representative days should be included in more realistic studies of power systems with large amounts of renewable [36].
3) In the case studies, multiple economic factors can affect the optimized planning portfolio, *e.g.*, the increasing fuel cost, decreasing renewable investment cost, annual interest rate, and the salvage income. Different economic settings may lead to a different result, but the conclusion of the renewable-involved G&TEP study should hold similarly as discussed in the paper. Besides, when we consider a longer term of planning, renewables will begin to show the potential of higher cost-effectiveness in the G&TEP.
4) Other practical factors should be taken into account in future research to make more precise investment decisions for investors, including but not limited to the employment of energy storage devices, more types of renewables, demand response, and ancillary services.

This paper aims at providing a general framework and a solution algorithm, which can be applied to different uncertainty-based G&TEP studies. The case studies in this paper are carried out under simplified simulation settings and demonstrate different renewable investment scenarios under specific parameters. The observations of a more complete and comprehensive long-term G&TEP study for realistic large-scale power systems may vary from this paper as the findings are dependent upon many sensitive parameters including cost, discount rate and other detailed operational constraints.

VII. Conclusion

This paper introduces a novel modeling and solution strategy for the generation and transmission expansion planning under ultra-high renewable penetration, which allows analyses of discrete and continuous uncertainties. Based on the theoretical derivation and numerical experiments, several remarks are in order:

- We propose a general hybrid stochastic and robust model that can accurately capture the uncertainties in the modern power grid, whose discrete and continuous features are taken into consideration with high flexibility.
- We propose a combinatorial solution strategy leveraging the state-of-the-art C&CG and L-shaped algorithms that can efficiently tackle intractable and multiplex uncertainty-based planning problems.
- We investigate the long-term renewable cost-effectiveness in the test results. A proper portfolio of generation mix of the conventional and renewable generation is shown to be beneficial under our specific problem settings, and we also pave a way for future discussions on the 100% renewable penetration potentials by investigating the long-term cost-effectiveness of the renewable generation.

Appendix A
Linear Relaxation Technique

The product q of one binary variable z and one continuous variable x can be relaxed as follows. The relaxation is tight due to the convexity of the McCormick Envelope.

$$q \leq x + M(1 - z)$$
$$q \geq x - M(1 - z)$$

$$q \leq Mz$$
$$q \geq -Mz$$

References

[1] "Examination of federal financial assistance in the renewable energy market: Implications and opportunities for commercial deployment of small modular reactors," Tech. Rep., Scully Capital, 2018.

[2] S. Hong, H. Cheng, and P. Zeng, "N-k constrained composite generation and transmission expansion planning with interval load," *IEEE Access*, vol. 5, pp. 2779–2789, 2017.

[3] L. Baringo and A. Baringo, "A stochastic adaptive robust optimization approach for the generation and transmission expansion planning," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 792–802, Jan. 2018.

[4] A. Moreira, D. Pozo, A. Street, and E. Sauma, "Reliable renewable generation and transmission expansion planning: Co-optimizing system's resources for meeting renewable targets," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3246–3257, Jul. 2017.

[5] A. Baharvandi, J. Aghaei, T. Niknam, M. Shafie-Khah, R. Godina, and J. P. S. Catalao, "Bundled generation and transmission planning under demand and wind generation uncertainty based on a combination of robust and stochastic optimization," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1477–1486, Jul. 2018.

[6] Q. Wang, J. Watson, and Y. Guan, "Two-stage robust optimization for N-k contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2366–2375, Aug. 2013.

[7] A. Moreira, A. Street, and J. M. Arroyo, "An adjustable robust optimization approach for contingency-constrained transmission expansion planning," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2013–2022, Jul. 2015.

[8] B. Chen, J. Wang, L. Wang, Y. He, and Z. Wang, "Robust optimization for transmission expansion planning: Minimax cost vs. minimax regret," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3069–3077, Nov. 2014.

[9] R. A. Jabr, "Robust transmission network expansion planning with uncertain renewable generation and loads," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4558–4567, Nov. 2013.

[10] F. Beltrán, W. de Oliveira, and E. C. Finardi, "Application of scenario tree reduction via quadratic process to medium-term hydrothermal scheduling problem," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4351–4361, Nov. 2017.

[11] X. Zhang and A. J. Conejo, "Robust transmission expansion planning representing long- and short-term uncertainty," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1329–1338, Mar. 2018.

[12] L. A. Gallego, L. P. Garces, M. Rahmani, and R. A. Romero, "High-performance hybrid genetic algorithm to solve transmission network expansion planning," *Inst. Eng. Technol. Gener., Transmiss. Distrib.*, vol. 11, no. 5, pp. 1111–1118, 2017.

[13] C. Roldán, R. Mínguez, R. García-Bertrand, and J. M. Arroyo, "Robust transmission network expansion planning under correlated uncertainty," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2071–2082, May 2019.

[14] E. Du *et al.*, "The role of concentrating solar power toward high renewable energy penetrated power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6630–6641, Nov. 2018.

[15] Z. Zhuo *et al.*, "Transmission expansion planning test system for AC/DC hybrid grid with high variable renewable energy penetration," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2597–2608, Jul. 2020.

[16] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Oper. Res. Lett.*, vol. 41, pp. 457–461, 2013.

[17] R. M. Van Slyke and R. Wets, "L-shaped linear programs with applications to optimal control and stochastic programming," *SIAM J. Appl. Math.*, vol. 17, no. 4, pp. 638–663, 1969.

[18] Q. Wang, J. Watson, and Y. Guan, "Two-stage robust optimization for N–K contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2366–2375, Aug. 2013.

[19] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.

[20] "2013 Assessment of Reliability Performance," Tech. Rep., Texas Reliability Entity, Apr. 2014.

[21] IIT Power Group, [Online]. Available: http://motor.ece.iit.edu/data/, *Index of Data*, 2018.

[22] GAMS Development Corporation, Washington, DC, USA, *GAMS - A User's Guide, GAMS Release 25.3.0*, 2018.

[23] M. R. Bussieck, M. C. Ferris, and T. Lohmann, *GUSS: Solving Collections of Data Related Models Within GAMS*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 35–56, 2012.

[24] "Capital cost estimates for utility scale electricity generating plants," Tech. Rep., U.S. Energy Information Administrative, 2016.

[25] "Cost and performance characteristics of new generating technologies, annual energy outlook 2020," Tech. Rep., U.S. Energy Information Administrative, 2020.

[26] F. Deng, X. Zeng, and L. Pan, "Research on multi-terminal traveling wave fault location method in complicated networks based on cloud computing platform," *Protection Control Modern Power Syst.*, vol. 2, May 2017.

[27] ERCOT, [Online]. Available: http://www.ercot.com/gridinfo, *Generation & Load - Grid Information*.

[28] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming," *Math. Program.*, vol. 95, pp. 493–511, Mar. 2003.

[29] H. Heitsch and W. Römisch, "Scenario tree reduction for multistage stochastic programs," *Comput. Manage. Sci.*, vol. 6, pp. 117–133, May 2009.

[30] A. Vafamehr, M. E. Khodayar, S. D. Manshadi, I. Ahmad, and J. Lin, "A framework for expansion planning of data centers in electricity and data networks under uncertainty," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 305–316, Jan. 2019.

[31] M. Farhoumandi, F. Aminifar, and M. Shahidehpour, "Generation expansion planning considering the rehabilitation of aging generating units," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3384–3393, Jul. 2020.

[32] "HOMER pro user manual 3.13 - glossary," Tech. Rep., HOMER Energy, 2019.

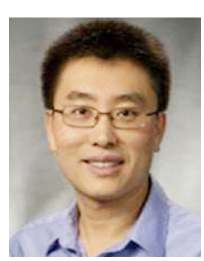
[33] J. Koornneef, T. van Keulen, A. Faaij, and W. Turkenburg, "Life cycle assessment of a pulverized coal power plant with post-combustion capture, transport and storage of CO_2," *Int. J. Greenhouse Gas Control*, vol. 2, no. 4, pp. 448–467.

[34] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, and I. E. Grossmann, "Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm," *Eur. J. Oper. Res.*, vol. 271, no. 3, pp. 1037–1054, 2018.

[35] J. E. Price, "Benchmarking a reduced test-bed model of WECC region for unit commitment and flexible dispatch," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2013, pp. 1–5.

[36] N. Helistö, J. Kiviluoma, H. Holttinen, J. D. Lara, and B.-M. Hodge, "Including operational aspects in the planning of power systems with large amounts of variable generation: A review of modeling approaches," *WIREs Energy Environ.*, vol. 8, no. 5, p. e341, 2019.

Shengfei Yin (Student Member, IEEE) received the B.Eng. degree from Hunan University, Changsha, China, in 2016, and the M.S. degree from the Illinois Institute of Technology, Chicago, IL, USA, in 2017, both in electrical engineering. He is currently working toward the Ph.D. degree with Southern Methodist University, Dallas, USA. His research interests include electricity market designs and the application of operations research and data analysis in the power system. He is a Reviewer for IEEE TRANSACTIONS ON SUSTAINABLE ENERGY and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, etc.



Jianhui Wang (Senior Member, IEEE) is a Professor with the Department of Electrical and Computer Engineering, Southern Methodist University. He has authored or coauthored more than 300 journal and conference publications, which have been cited for more than 20,000 times by his peers with an H-index of 80. He has been invited to give tutorials and keynote speeches at major conferences including IEEE ISGT, IEEE SmartGridComm, IEEE SEGE, IEEE HPSC and IGEC-XI.

Dr. Wang is the past Editor-in-Chief of the IEEE TRANSACTIONS ON SMART GRID and an IEEE PES Distinguished Lecturer. He is also a Guest Editor of a *Proceedings of the IEEE* Special Issue on power grid resilience. He was the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015 and the 2018 Premium Award for Best Paper in IET Cyber-Physical Systems: Theory & Applications. He is a 2018 and 2019 Clarivate Analytics highly cited researcher for production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science.