

# Mixed-integer Linear Programming Models and Algorithms for Generation and Transmission Expansion Planning of Power Systems

Can Li<sup>a</sup>, Antonio J. Conejo<sup>b,c</sup>, Peng Liu<sup>d</sup>, Benjamin P. Omell<sup>d</sup>, John D. Siirola<sup>e</sup>, Ignacio E. Grossmann<sup>a,\*</sup>

<sup>a</sup>*Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA*

<sup>b</sup>*Department of Integrated Systems Engineering, The Ohio State University, 1971 Neil Avenue, Columbus, OH 43210, USA*

<sup>c</sup>*Department of Electrical and Computer Engineering, The Ohio State University, 2015 Neil Avenue, Columbus, OH 43210, USA*

<sup>d</sup>*National Energy Technology Laboratory, Pittsburgh, PA 15236, United States*

<sup>e</sup>*Center for Computing Research, Sandia National Laboratories, P.O. 5800, Albuquerque, NM, 87185, USA*

---

## Abstract

With the increasing penetration of renewable generating units, especially in remote areas not well connected with load demand, there are growing interests to co-optimize generation and transmission expansion planning (GTEP) in power systems. Due to the volatility in renewable generation, a planner needs to include the operating decisions into the planning model to guarantee feasibility. However, solving the GTEP problem with hourly operating decisions throughout the planning horizon is computationally intractable. Therefore, we propose several spatial and temporal simplifications to the problem. Built on the generation expansion planning (GEP) formulation of Lara et al. (2018), we propose a mixed-integer linear programming formulation for the GTEP problem. Three different formulations, i.e., a big-M formulation, a hull formulation, and an alternative big-M formulation, are reported for transmission expansion. We theoretically compare the tightness of the LP relaxations of the three for-

---

\*Corresponding author

Email addresses: [canli@andrew.cmu.edu](mailto:canli@andrew.cmu.edu) (Can Li), [conejo.1@osu.edu](mailto:conejo.1@osu.edu) (Antonio J. Conejo), [Peng.Liu@netl.doe.gov](mailto:Peng.Liu@netl.doe.gov) (Peng Liu), [Benjamin.Omell@netl.doe.gov](mailto:Benjamin.Omell@netl.doe.gov) (Benjamin P. Omell), [jdsiiro@sandia.gov](mailto:jdsiiro@sandia.gov) (John D. Siirola), [grossmann@cmu.edu](mailto:grossmann@cmu.edu) (Ignacio E. Grossmann)

mulations. The proposed MILP GTEP model typically involves millions or tens of millions of variables, which makes the model not directly solvable by the commercial solvers. To address this computational challenge, we propose a nested Benders decomposition algorithm and a tailored Benders decomposition algorithm that exploit the structure of the GTEP problem. Using a case study from Electric Reliability Council of Texas (ERCOT), we are able to show that the proposed tailored Benders decomposition outperforms the nested Benders decomposition. The coordination in the optimal generation and transmission expansion decisions from the ERCOT study implies that there is an additional value in solving GEP and TEP simultaneously.

*Keywords:* OR in energy, Power Systems, Generation Transmission Expansion, Mixed-integer Programming, Decomposition Algorithm

---

## 1. Introduction

Generation expansion planning (GEP) of power systems involves determining the optimal size, location, and construction time of new power generation plants, while minimizing the total cost over a long-term planning horizon (Conejo et al., 2016; Koltsaklis and Dagoumas, 2018). There is a growing interest to use mathematical programming models to solve generation expansion planning problems (Lara et al., 2018; Sadeghi et al., 2017; Oree et al., 2017). Conventional power units are dispatchable thermal power plants that can provide stable power output. Due to computational tractability concerns, generation expansion models can ignore short-term operating decisions. However, with the increased penetration of renewable generation technologies, such as solar and wind, power systems nowadays need to be more flexible so as to adjust to the volatile power generation from renewables. In this case, operations decisions, such as unit commitment, ramping decisions, become important to assess system feasibility (Ding and Somani, 2010; Koltsaklis and Georgiadis, 2015; Pina et al., 2013; Poncelet et al., 2014; Shortt and O’Malley, 2010; Flores-Quiroz et al., 2016; Palmintier and Webster, 2011; Lara et al., 2018; Lohmann

and Rebennack, 2017). Due to the incorporation of short-term operating constraints into the long-term planning problem, the integrated model is computationally challenging. In order to solve such multi-scale problem efficiently, Lara et al. (2018) use nested Benders decomposition to solve a GEP model with unit commitment. Lohmann and Rebennack (Lohmann and Rebennack, 2017) develop a tailored Generalized Benders Decomposition algorithm.

Transmission expansion planning (TEP) refers to installing new transmission lines or expanding the capacities of existing transmission lines in a power system. Bahiense et al. (2001) propose a mixed integer disjunctive model for transmission network expansion. (Alguacil et al., 2003) propose an MILP model that considers losses and guarantees convergence to optimality for the TEP. Zhang et al. (2013) propose an improved model that includes a linear representation of reactive power, off-nominal bus voltage magnitudes and network losses. For a more detailed review of of TEP models and algorithms, we refer the readers to the review papers (Hemmati et al., 2013; Ude et al., 2019).

GEP and TEP are generally solved as two independent optimization problems since the market agents addressing these two problems are different. GEP pertains to producers, while TEP pertains to a regulated planner. However, the significant penetration of renewables into power systems may lead to their concentration in remote areas not well connected to load demand (Koltsaklis and Dagoumas, 2018). Therefore, installing renewables in those remote areas could compromise transmission expansion. The recognition of transmission's interaction with generation expansion has motivated the development of co-optimization methods to consider the tradeoffs between generation and transmission expansion (Krishnan et al., 2016). Several works have been reported to simultaneously optimize generation and transmission expansion planning (GTEP) (Pozo et al., 2012; Aghaei et al., 2014). We refer to Table 1 of the review paper (Koltsaklis and Dagoumas, 2018) for a long list of works. See also the review paper (Gacitua et al., 2018).

A number of related works consider uncertainties in the planning problem using two-stage or multistage stochastic programming (Lara et al., 2019; O'Neill

et al., 2013; Liu et al., 2017), robust optimization (Mejía-Giraldo and McCalley, 2013; Baringo and Baringo, 2017). (Le Cadre et al., 2015; Pozo et al., 2012) apply game theory or multi-level optimization to characterize the interaction of the participants in the markets.

This paper is an extension of the GEP model reported in Lara et al. (2018) to a GTEP model. In Lara et al. (2018), the authors propose an MILP model for deterministic generation expansion planning problem that represents the hourly operating decisions of the generators and storage units. Renewable generation and load data on some representative days are used as the input to the hourly unit commitment model (Mallapragada et al., 2018). Lara et al. (2018) use a tailored nested Benders decomposition algorithm to solve the multi-scale GEP problem. However, in Lara et al. (2018) transmission expansion planning is not considered and the power flow equations ignore Kirchhoff's voltage law.

The major contributions of this paper are listed below.

- We extend the model in Lara et al. (2018) by considering transmission expansion and DC power flow equations.
- Three different formulations for transmission expansion, i.e., big-M formulation, hull formulation, and an alternative big-M formulation proposed by Bahiense et al. (2001) are investigated.
- The proposed GTEP model is computationally more challenging to solve than the GEP model in Lara et al. (2018). Regarding solution technique, the novel contribution of this paper is a tailored Benders decomposition algorithm to solve the GTEP problem. We compare the nested Benders algorithm (Lara et al., 2018) and the tailored Benders decomposition algorithm for the new GTEP model.
- The case study demonstrates the importance of the coordination between the generation and the transmission decisions in the optimal solution.

The rest of this paper is organized as follows. In section 2, we give the description and the assumptions of the problem that we address. In section 3, we describe

the MILP formulation for our GTEP model. In section 4, we describe two solution techniques, a nested Benders decomposition and a tailored Benders decomposition. In section 5, a case study from Electric Reliability Council of Texas (ERCOT) is used to illustrate the working of the model and the efficiency of the solution techniques. We draw the conclusion in section 6

## 2. Problem Statement and Assumptions

Given is a geographical region with existing and potential generating units and transmissions lines. The problem consists in making capacity expansion decisions for both generation and transmission while considering the unit commitment and power flow constraints at the operational level.

### 2.1. Generation representation

The existing and potential generation technologies are similar to the ones used in Lara et al. (2018), i.e.,

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

Also known are: the generating units' nameplate (maximum) capacity; expected lifetime; fixed and variable operating costs; fixed and variable start-up cost; cost for extending their lifetimes; CO<sub>2</sub> emission factor and carbon tax, if applicable;

fuel price, if applicable; and operating characteristics such as ramp-up/ramp-down rates, operating limits, contribution to spinning and quick start fraction for thermal generators, and capacity factor for renewable generators.

For the case of existing generators, their age at the beginning of the study horizon and location are also known. For the case of potential generators, the capital cost and the maximum yearly installation of each generation technology are also given. Also given is a set of potential storage units, with specified technology (e.g., lithium ion, lead-acid, and flow batteries), capital cost, power rating, rated energy capacity, charge and discharge efficiency, and storage lifetime. Additionally, the projected load demand is given for each location.

We assume that the generators using the same type of technology are homogeneous, i.e., their design parameters are identical. For example, all the coal-st-old generators have the same parameters, which can be obtained by performing aggregation on the existing generators that use coal steam turbines. Note that although the renewable generators of the same technology have the same design parameters under our assumption, they can have different capacity factors depending on the weather conditions of the region in which they are installed.

## *2.2. Transmission representation*

Given are existing and candidate transmission lines between any of the two neighboring buses. The susceptance, distance, and capacity of each transmission line are known. For the existing transmission lines, we assume that they will not reach their life expectancy during the planning horizon, i.e., we do not consider the retirement of transmission lines. For the candidate transmission lines, the capital cost of each transmission line is known.

We use DC power flow equations to calculate the power flow in each transmission line. These equations are built based on Kirchhoff's voltage and current laws which differ from the network flow model used in the work of Lara et al. (2018). In the network flow model, the transmission network is represented similarly to pipelines where the flows only observe energy balance at each node

while ignoring Kirchhoff's laws.

### *2.3. Temporal representation*

The GTEP model integrates unit commitment decisions to evaluate the hourly operation requirements. Given that the planning horizon of the GTEP problem can be as long as 10 to 30 years, solving the long-term planning problem with operating decisions in every hour of the planning horizon is intractable. Therefore, a simplification is needed to make the problem solvable, while representing the hourly fluctuations of the load and renewable profiles.

Several works propose to select a few representative days (Mallapragada et al., 2018; Teichgraeber and Brandt, 2019; Scott et al., 2019) from the full data set to represent the hourly fluctuations. To keep the chronology of the hourly historical data, the time series for the loads and the capacity factors corresponding to the same day are concatenated as a single vector, which will be used as the input to some clustering algorithms, such as  $k$ -means, and hierarchical clustering. After performing the clustering on the full time series data set, the time series corresponding to the representative days are the centroids or medoids of the clusters. The details can be found in subsection 5.1.

### *2.4. Spatial representation*

GTEP is typically performed on large scale power systems which consists of thousands of buses, such as ERCOT, SPP, PJM, MISO, etc. In most cases, it is intractable for GTEP to model each bus. Therefore, we adopt a similar approach as in Lara et al. (2018) to reduce the spatial complexity of the problem. The area of interest is divided into several regions that have similar climate (e.g., wind speed and solar incidence over time), and load profiles. As we describe in the generation representation subsection, all the generators using the same technology have the same parameters. On the other hand, for the renewable generators, the capacity factors are dependent on the location at which they are installed. We assume that the capacity factors of the renewable generators in the same region are the same.

We assume that all the generators and loads are located at the center of each region. Since each region is treated as one bus in the power flow model, we only consider the tielines between two neighboring regions. We assume that the two ends of each tieline are the centers of the two regions it connects. All the tielines are assumed to have the same voltage, susceptance, and capacity. An example of the proposed spatial representation approach is shown in Figure 1. The ERCOT region is divided into five regions, Panhandle, Northeast, West, South, and Coast. The center of each region is specified as one of the cities in the region. The existing transmission lines are represented as solid lines while the candidate transmission lines are represented as dashed lines. Each region has generator clusters corresponding to different technologies.

The aggregation of the generating units is a simplification of the problem that may yield suboptimal solution compared with modeling each generator individually. Such simplification is necessary to make the problem tractable. In order to obtain a feasible solution to the real physical system, i.e., the unit commitment decisions of each generator, one could perform a disaggregation heuristics on the aggregated solution. We will leave developing these heuristics as future work.

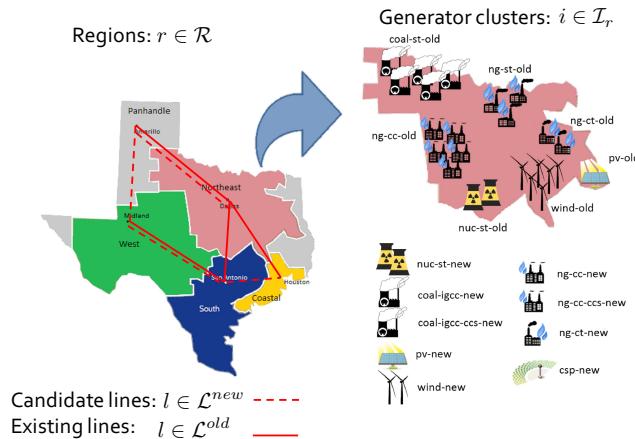


Figure 1: Spatial representation of the five ERCOT regions' generator clusters and transmission lines

## 2.5. Decisions and objective

With the above input data, spatial and temporal representations, the proposed GTEP model is to decide: a) when and where to install new generators, storage units and transmission lines; b) when to retire generators and storage units; c) whether or not to extend the life of the generators that reached their expected lifetime; d) unit commitment of the thermal generators during the representative days; e) power generations of the generator clusters and power flows through the transmission lines. The objective is to minimize the overall cost including operating, investment, and environmental costs (e.g., carbon tax and renewable generation quota).

## 3. MILP Formulation

This section presents a deterministic MILP formulation for the GTEP problem. Most of the MILP formulation is similar to that in Lara et al. (2018). Here, we emphasize the transmission expansion formulation that is added. Note that if an index appears in a summation or next to a  $\forall$  symbol without a set, all elements in the corresponding set should be considered. The nomenclature for sets, parameters, and variables used in the MILP formulation are provided in Appendix A in supplementary material.

### 3.1. Transmission expansion constraints

Transmission line balance constraints. A succinct version of Appendix A is attached at the end of this paper. Variable  $ntb_{l,t}$  denotes whether or not candidate transmission line  $l$  is built in year  $t$ . Variable  $nte_{l,t}$  denotes whether transmission line  $l$  has been installed in year  $t$ . Equation (1) represents the balance of transmission lines.

$$nte_{l,t} = nte_{l,t-1} + ntb_{l,t} \quad \forall l \in \mathcal{L}^{\text{new}}, t \quad (1)$$

The DC transmission constraints calculate and limit the power flows through the transmission lines. Parameter  $B_l$  represents the susceptance of line  $l$ .

$\theta_{s(l),t,d,s}$ ,  $\theta_{r(l),t,d,s}$  are the phase angles of the buses that are the sending-end and the receiving-end of line  $l$ , respectively, in year  $t$ , representative day  $d$ , and sub-period (hour)  $s$ . The existing transmission lines have to satisfy the DC power flow equation (2).

$$p_{l,t,d,s}^{\text{flow}} = B_l(\theta_{s(l),t,d,s} - \theta_{r(l),t,d,s}) \quad \forall l \in \mathcal{L}^{\text{old}}, t, d, s \quad (2)$$

It should be noted that the DC power flow equation is an approximation of the AC power flow equation. The approximation is accurate under some assumptions (Frank and Rebennack, 2016): all system branch resistances are approximately zero; the differences between adjacent bus voltage angles are small; the system bus voltages are approximately equal to the 1.0 per unit; reactive power flow is neglected.

The power flow through each transmission line is bounded. Parameter  $F_l^{\max}$  represents the capacity of transmission line  $l$ . Thus:

$$-F_l^{\max} \leq p_{l,t,d,s}^{\text{flow}} \leq F_l^{\max} \quad \forall l \in \mathcal{L}^{\text{old}}, t, d, s \quad (3)$$

For the candidate transmission lines, we can write the following disjunction, where  $NTE_{l,t}$  is a logic variable whose value can be True or False indicating whether or not transmission line  $l$  is installed in year  $t$ . If line  $l$  already exists in year  $t$ , the corresponding power flow has to satisfy DC power flow equation and upper and lower bounds. Otherwise, the corresponding power flow is zero. We assume that all the candidate transmission lines are standard. In other words, the susceptance of the candidate transmission lines  $B_l$  are parameters in the model.

$$\left[ \begin{array}{c} NTE_{l,t} \\ p_{l,t,d,s}^{\text{flow}} = B_l(\theta_{s(l),t,d,s} - \theta_{r(l),t,d,s}) \\ -F_l^{\max} \leq p_{l,t,d,s}^{\text{flow}} \leq F_l^{\max} \end{array} \right] \vee \left[ \begin{array}{c} \neg NTE_{l,t} \\ p_{l,t,d,s}^{\text{flow}} = 0 \end{array} \right] \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (4)$$

Standard approaches, i.e., big-M reformulation and hull reformulation (Gross-

mann and Trespalacios, 2013), are available to reformulate disjunctions (4) into mixed integer constraints.

The big-M formulation of the disjunction is,

$$-M_l(1-n_{te_{l,t}}) \leq p_{l,t,d,s}^{\text{flow}} - B_l(\theta_{s(l),t,d,s} - \theta_{r(l),t,d,s}) \leq M_l(1-n_{te_{l,t}}) \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (5)$$

$$-F_l^{\max} n_{te_{l,t}} \leq p_{l,t,d,s}^{\text{flow}} \leq F_l^{\max} n_{te_{l,t}} \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (6)$$

This big-M formulation is most commonly used in the literature (Conejo et al., 2016) for TEP.

The hull formulation is,

$$p_{l,t,d,s}^{\text{flow}} = B_l \Delta\theta_{l,t,d,s}^1 \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (7)$$

$$\theta_{s(l),t,d,s} - \theta_{r(l),t,d,s} = \Delta\theta_{l,t,d,s}^1 + \Delta\theta_{l,t,d,s}^2 \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (8)$$

$$-\pi \cdot n_{te_{l,t}} \leq \Delta\theta_{l,t,d,s}^1 \leq \pi \cdot n_{te_{l,t}} \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (9)$$

$$-\pi(1 - n_{te_{l,t}}) \leq \Delta\theta_{l,t,d,s}^2 \leq \pi(1 - n_{te_{l,t}}) \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (10)$$

where  $\Delta\theta_{l,t,d,s}^1$  and  $\Delta\theta_{l,t,d,s}^2$  are disaggregated variables for the angle difference of transmission line  $l$ . Variable  $\Delta\theta_{l,t,d,s}^1$  is equal to the angle difference if transmission line  $l$  has been installed in year  $t$ . Otherwise,  $\Delta\theta_{l,t,d,s}^2$  equals to the angle difference. In addition to equations (7)-(10), equation (6) needs to be included in the hull formulation.

The hull formulation has more continuous variables than the big-M formulation but it avoids using the big-M parameters of equations (5). The hull formulation provides the intersection of the convex hull of each disjunction in (4). Therefore, the hull formulation can provide a tighter LP relaxation at the

expense of solving larger LPs at each node of a branch-and-bound algorithm.

Alternative big-M formulation: Besides the big-M and hull formulations, an alternative big-M formulation is proposed by Bahiense et al. (2001). In this formulation, additional continuous variables  $p_{l,t,d,s}^{\text{flow+}}$ ,  $p_{l,t,d,s}^{\text{flow-}}$ ,  $\Delta\theta_{l,t,d,s}^+$ ,  $\Delta\theta_{l,t,d,s}^-$ , are introduced, where the superscript ‘+’ means that the flow is in the same direction as the nominal direction of transmission line  $l$ , i.e., from the sending-end node  $s(l)$  to the receiving-end node  $r(l)$ ; superscript ‘-’ means the opposite direction. By defining these new continuous variables, equation (5) is replaced by equations (11) to (14) and equation (6) is replaced by equations (17) and (18). Bahiense et al. (2001) claim that the alternative big-M formulation is tighter than the big-M formulation. However, we prove that they have the same feasible region if we project the feasible region of the alternative big-M formulation onto the space of  $(p_{l,t,d,s}^{\text{flow}}, \theta_{s(l),t,d,s}, \theta_{r(l),t,d,s}, nte_{l,t})$  in Theorem 1. The proof of Theorem 1 can be found in Appendix B of supplementary material

$$p_{l,t,d,s}^{\text{flow+}} - B_l \Delta\theta_{l,t,d,s}^+ \leq 0 \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (11)$$

$$p_{l,t,d,s}^{\text{flow-}} - B_l \Delta\theta_{l,t,d,s}^- \leq 0 \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (12)$$

$$p_{l,t,d,s}^{\text{flow+}} - B_l \Delta\theta_{l,t,d,s}^+ \geq -M_l(1 - nte_{l,t}) \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (13)$$

$$p_{l,t,d,s}^{\text{flow-}} - B_l \Delta\theta_{l,t,d,s}^- \geq -M_l(1 - nte_{l,t}) \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (14)$$

$$p_{l,t,d,s}^{\text{flow}} = p_{l,t,d,s}^{\text{flow+}} - p_{l,t,d,s}^{\text{flow-}} \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (15)$$

$$\theta_{s(l),t,d,s} - \theta_{r(l),t,d,s} = \Delta\theta_{l,t,d,s}^+ - \Delta\theta_{l,t,d,s}^- \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (16)$$

$$p_{l,t,d,s}^{\text{flow+}} \leq F_l^{\max} nte_{l,t} \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (17)$$

$$p_{l,t,d,s}^{\text{flow-}} \leq F_l^{\max} nte_{l,t} \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (18)$$

$$p_{l,t,d,s}^{\text{flow+}}, p_{l,t,d,s}^{\text{flow-}}, \Delta\theta_{l,t,d,s}^+, \Delta\theta_{l,t,d,s}^- \geq 0 \quad \forall l \in \mathcal{L}^{\text{new}}, t, d, s \quad (19)$$

**Theorem 1.** *The alternative big-M formulation (ABM) has the same feasible region as the big-M (BM) formulation if the feasible region of ABM is projected to the space of  $\left\{ \bigoplus_{l \in \mathcal{L}^{\text{new}}, t \in \mathcal{T}, d \in \mathcal{D}, s \in \mathcal{S}} (p_{l,t,d,s}^{\text{flow}}, \theta_{s(l),t,d,s}, \theta_{r(l),t,d,s}, nte_{l,t}) \right\}$ , where the symbol ‘ $\oplus$ ’ means the concatenation of all the variables  $(p_{l,t,d,s}^{\text{flow}}, \theta_{s(l),t,d,s}, \theta_{r(l),t,d,s}, nte_{l,t})$  over the set  $\mathcal{L}^{\text{new}}, \mathcal{T}, \mathcal{D}, \mathcal{S}$ .*

### 3.2. Other constraints

All other constraints including operational constraints, investment-related constraints, generator balance constraints, storage constraints, are similar to those of the MILP formulation proposed by Lara et al. (2018). The details of these constraints and the nomenclature can be found in Appendix A in supplementary material. A succinct version of Appendix A is attached at the end of this paper.

## 4. Solution techniques

Given the large size of the proposed GTEP problem, tailored solution approaches need to be developed. In this section, we describe two solution algorithms: a) nested Benders decomposition adapted from (Birge, 1985; Zou et al., 2019), which has been used by Lara et al. (2018) to solve the GEP model. b) a tailored Benders decomposition. Both of the two algorithms exploit the structure of the GTEP problem.

### 4.1. Nested Benders decomposition

Lara et al. (2018) apply a nested Benders decomposition algorithm to solve their GEP model. Like in the GEP model, the nested Benders decomposition algorithm decomposes the fullspace of the GTEP problem by year. Note

that the linking constraints for two consecutive years are the investment related constraints. For the investment decisions in transmission lines, the linking constraints are described by equation (1), i.e., the balance of candidate transmission lines. Similarly, there are linking constraints corresponding to the number of thermal generators  $ngo_{i,r,t}^{\text{th}}$ , the number of renewable generators  $ngo_{i,r,t}^{\text{rn}}$ , the number of storage units  $nso_{j,r,t}$  per region  $r$  and year  $t$ . These linking constraints can be found in equations (1), and (A.17), (A.20), (A.24) in Appendix A in supplementary material.

From the above observation, variables  $nte_{l,t}$ ,  $ngo_{i,r,t}^{\text{th}}$ ,  $ngo_{i,r,t}^{\text{rn}}$ ,  $nso_{j,r,t}$  can be treated as complicating variables. Once these variables are fixed, the GTEP problem can be decomposed by year. The nested Benders decomposition consists of two phases, i.e., forward pass and backward pass. In the forward pass, the problem is solved sequentially year after year. In each year  $t$ , the problem is solved in a myopic way, with the complicating variables of year  $t - 1$  fixed, and the cutting planes generated from the backward pass. The optimal solution is obtained for year  $t$ . Then the complicating variables are fixed for year  $t$  and the problem for year  $t + 1$  is solved, until we reach the end of the planning horizon.

In the backward pass, cutting planes can be generated by solving the LP relaxations of the planning problem with the complicating variables fixed at the values of the forward pass. The backward pass starts from the last year and sequentially adds cutting planes to the previous year. Since the nested Benders decomposition is developed by Lara et al. (2018) for the GEP model, we do not provide the details of the algorithms. The steps of the nested Benders algorithms are similar to those in Lara et al. (2018), except that in the GTEP problem we introduce new complicating variables  $nte_{l,t}$  pertaining to transmission expansion. An additional difference is that while in Lara et al. (2018) three different types of cutting planes are implemented in the backward pass, i.e., Benders cuts, strengthened Benders cuts, and Lagrangean cuts, we only implement Benders cuts to solve the GTEP problem because this type of cut is computationally cheap.

#### 4.2. Tailored Benders decomposition algorithm

Instead of solving the GTEP problem sequentially by year as in the nested Benders decomposition, we treat all the investment-related variables as complicating variables and include all these variables in a single Benders master problem.

More specifically, the proposed GTEP model can be represented using the succinct form (20) below, where  $x_t$  represents all the investment decisions in year  $t$ ,  $y_t$  represents all the operating decisions in the representative days for year  $t$ . Note that the investment decisions are made on a yearly basis indexed by  $t$ . The operating decisions not only have the index  $t$  but also have indices  $d$  and  $s$ , which represent the  $d$ th representative day in the  $s$ th hour, respectively. Since we will decompose the problem by year, we omit the indices  $d$  and  $s$  and simply use  $y_t$  to represent all the operating decisions corresponding to year  $t$ . Equations (20c) and (20d) are investment related constraints, which correspond to equations (1), (A.14)-(A.21), (A.23), (A.24). Equations in (20b) describe the operational decisions of each year, such as the power flow equations (2) and (3). Note that equation (20b) can be decomposed by year. Equation (20e) represents the integrality constraints and variable bounds that  $x_t$  and  $y_t$  have to satisfy.

$$\min \sum_{t \in \mathcal{T}} c_t^\top x_t + d_t^\top y_t \quad (20a)$$

$$\text{s.t. } A_t x_t + B_t y_t \leq b_t \quad \forall t \in \mathcal{T} \quad (20b)$$

$$C_1 x_1 \leq f_1 \quad (20c)$$

$$C_{t-1} x_{t-1} + D_t x_t \leq f_t \quad t = 2, 3, \dots, |\mathcal{T}| \quad (20d)$$

$$x_t \in X_t, y_t \in Y_t \quad \forall t \in \mathcal{T} \quad (20e)$$

The GTEP problem has a decomposable structure in the sense that if we treat all the investment decisions  $x_t$  for all  $t \in \mathcal{T}$  as *complicating variables*, the problem can be decomposed by year. Benders decomposition (Rahmaniani et al., 2016)

can be applied to solve this type of problem. We can assign all the investment variables to the Benders master problem and the operating variables  $y_t$  to the  $t$ th subproblem. After solving the Benders master problem, the investment decisions are fixed and each Benders subproblem can be solved independently. Note that there are some integer variables in the operating decisions, such as the number of generators that are on/off. In order to generate valid Benders cuts, we solve the LP relaxation of each Benders subproblem and add the cuts to the Benders master problem. A high level description of the algorithm is provided in Figure 2. The formulation of the Benders master problem solved at

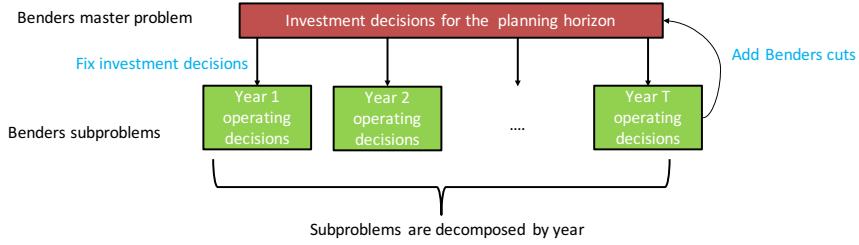


Figure 2: Tailored Benders decomposition algorithm applied to the GTEP problem

iteration  $k$  is:

$$\min \sum_{t \in \mathcal{T}} c_t^\top x_t + \eta_t \quad (21a)$$

$$\text{s.t. } C_1 x_1 \leq f_1 \quad (21b)$$

$$C_{t-1} x_{t-1} + D_t x_t \leq f_t \quad t = 2, 3, \dots, |\mathcal{T}| \quad (21c)$$

$$\eta_t \geq \tilde{\eta}_t^{k'} + (\mu_t^{k'})^\top (\tilde{x}_t^{k'} - x_t) \quad t \in \mathcal{T}, k' < k \quad (21d)$$

$$x_t \in X_t \quad \forall t \in \mathcal{T} \quad (21e)$$

where equation (21d) are the Benders cuts generated by solving the Benders subproblems. We denote the optimal solution of the Benders master problem at iteration  $k$  as  $\tilde{x}_t^k, \forall t \in \mathcal{T}$ .

Fixing the values of the investment decision variables to the values obtained

at the master problem, i.e.,  $x_t = \tilde{x}_t^k, \forall t \in \mathcal{T}$ , we can solve each Benders subproblem independently for each year  $t \in \mathcal{T}$ :

$$\tilde{\eta}_t^k = \min d_t^\top y_t \quad (22a)$$

$$\text{s.t. } x_t = \tilde{x}_t^k \quad (22b)$$

$$A_t x_t + B_t y_t \leq b_t \quad (22c)$$

$$y_t \in \tilde{Y}_t \quad (22d)$$

where all the integer variables in  $y_t$  are relaxed and set  $\tilde{Y}_t$  represents set  $Y_t$  without the integrality constraints, i.e.,  $\tilde{Y}_t$  only represents variables bounds for  $y_t$ . Let  $\mu_t^k$  be the optimal dual multiplier for equation (22b). A valid Benders cut,

$$\eta_t \geq \tilde{\eta}_t^k + (\mu_t^k)^\top (\tilde{x}_t^k - x_t)$$

can be generated by solving the  $t$ th subproblem. The cuts from the subproblems are then added to the master problem via equation (21d). Note that the Benders subproblem (22) can be infeasible. In this case, a feasibility subproblem should be solved to generate feasibility cuts. Interested readers can refer to (Rahmaniani et al., 2016) for the definitions of feasibility cuts. To simply notation, we assume that the subproblems are feasible here.

At each iteration  $k$ , the Benders master problem provides a lower bound of the optimal objective function value with relaxed  $y_t$  variables, while an upper bound can be calculated as  $\sum_{t \in \mathcal{T}} c_t^\top \tilde{x}_t^k + d_t^\top \tilde{y}_t^k$  where  $\tilde{x}_t^k$  and  $\tilde{y}_t^k$  are the optimal solutions to the master problem and the subproblems, respectively. We keep iterating between the Benders master problem (21) and the subproblems (22) until the upper bound and the lower bound lie within certain optimality tolerance.

For our computational study, we use the Benders implementation from

CPLEX (Bonami et al., 2020), which is a branch-and-Benders-cut algorithm. We only need to specify the variables in the master problem and the variables in each subproblem and CPLEX automatically solves the GTEP problem using Benders decomposition. Note that the implementation in CPLEX uses a single branch-and-bound tree where Benders cuts are added as lazy constraints. The Benders subproblems are solved whenever a feasible solution is found in the branch-and-bound tree of the master problem. The corresponding optimality or feasibility cuts will be added to the master problem dynamically in the single branch-and-bound tree.

Since the integrality constraints of the  $y_t$  variables are relaxed within the Benders decomposition algorithm, we can only obtain a lower bound to the original GTEP problem (20) through this algorithm. In order to obtain a feasible solution to the original problem, i.e., an upper bound, we can fix the investments decisions  $x_t$  to the optimal solution of the Benders decomposition algorithm and solve the operating problem with the integrality constraints of the  $y_t$  variables for each year independently. Moreover, as a result of the relaxation of the integrality constraints corresponding to the  $y_t$  variables, there will be a gap between the lower bound and the upper bound. However, our computational results in section 5 show that this gap is small. The reason is that all the integer variables in  $y_t$  are general integer variables instead of binary variables. Typically, mixed-integer programs with general integer variables have good LP relaxations.

## 5. Case studies

### 5.1. Input data

We carry out a GTEP case study for ERCOT. The spatial representation of the ERCOT region has been discussed in subsection 2.4. It is divided into four geographical regions: Northeast, West, Coast and South. Besides these regions, a fifth region, Panhandle, is also included, which is technically outside the ERCOT region, but due to its renewable generation potential, it supplies

electricity to the ERCOT regions. Note that in our model, Panhandle is treated as a pure supplier, i.e., it has zero load. The map of the five regions is shown in Figure 1.

Each of the five regions are treated as a bus and a DC power flow model is considered. We specify a city for each region as the location of the bus. The center for Northeast, West, Coast, South and Panhandle are Dallas, Midland, Houston, San Antonio, and Amarillo, respectively. The lengths of the transmission lines are determined by the distance between the centers of any of the two neighboring regions. To test the GTEP model, we assume that no transmission lines are available, i.e., the model will identify the transmission lines to be built. We assume that for each pair of the two neighboring regions, at most 10 candidate transmission lines can be built. The susceptance and capacity of the transmission lines are all the same, which are obtained from a synthetic grid of Texas (Birchfield et al., 2016). The unit capital cost of transmission lines is \$1,919,450 per mile obtained from (Andrade and Baldick, 2016).

Old and new generation technologies have been described in subsection 2.1. The investment cost, fixed and variable operating costs for different generation technologies are obtained from the National Renewable Energy Laboratory (NREL), available in the 2016 Annual Technology Baseline (ATB) Spreadsheet (Cole et al., 2016). The capital cost, power rating, rated energy capacity, charge and discharge efficiency and storage lifetime of the storage units are from (Schmidt et al., 2017). We consider a 20 year time horizon, in which the first year is 2019. The fuel price data for coal, natural gas and uranium correspond to the reference scenario in U.S. Energy Information Administration (2019). A discount rate of 5.7% as chosen in Short et al. (2011) is used. We assume that the undiscounted carbon tax is zero in the first year and grows linearly across years to \$325/tonne CO<sub>2</sub>, which is on the high side compared with most forecast scenarios reported in McFarland et al. (2018). The curtailment cost is assumed to be \$5,000/MWh.

The hourly solar capacity factor profiles including photo-voltaic (pv) and concentrated solar power (csp), are calculated based on the national solar radi-

ation data base (NSRDB) (Sengupta et al., 2018) in 2012 via the System Advisor Model (SAM) (Blair et al., 2014). The hourly wind capacity factor profiles are calculated based on the wind speed from the wind integration national dataset (wind) toolkit (Draxl et al., 2015) in 2012 using one power curve from SAM. Since load data are correlated with solar and wind capacity factors, to generate the hourly load profiles we take load data from ERCOT in 2012 and scale them so that the annual load for each ERCOT region is equal to the annual load in 2019. To sum up, we have 365 days (the leap day is excluded) of 24 hour solar and wind capacity factor and load data. The capacity factors are assumed to be unchanged over the planning horizon. The annual load growth rate is assumed to be 1.4% calculated based on the historical load data from 2011 to 2018 (ERCOT, 2106). To select the representative days for the GTEP model, we use the software package, TimeSeriesClustering.jl developed by Teichgraeber and Brandt (Teichgraeber and Brandt, 2019). By using this package, we are able to apply the  $k$ -means clustering algorithm to the time series and select the centroid of each cluster as the representative day. The weight of each representative day is proportional to the number of data points in that cluster. The details of the clustering algorithms are described in Teichgraeber and Brandt (2019). There is a trade-off between computational complexity and model fidelity in selecting the number of representative days. Here, we adopt a trial-and-error approach and gradually increase the number of representative days until the investment decisions do not change significantly. Due to the length constraint of the paper, the sensitivity analysis with respect to the number of representative days is given in Appendix D of the supplementary material. The results with 15 representative days are reported in subsection 5.3 because the solutions stabilize after the number of representative days is increased to 15.

### 5.2. Comparison of formulations and algorithms

All the MILP formulations are implemented in Pyomo/Python (Hart et al., 2011). We first solve the GTEP model directly with CPLEX 12.9.0.0. We compare the three transmission expansion formulations proposed in subsection 3.1.

All the problems in this paper are solved using one processor of an Intel Xeon (2.67GHz) machine with 64 GB RAM. The time limit is set to 10 hours. The number of general integer variables, binary variables, continuous variables, and constraints of the fullspace GTEP problem with the three different formulations are given in Table 1. The upper bound (UB), lower bound (LB) of the optimal value of the objective function in billion dollars and the wall time in seconds are also shown in Table 1. All the three formulations have the same number of general integer variables and binary variables but differ in the number of continuous variables. The standard big-M formulation uses the fewest number of continuous variables and constraints. CPLEX is not able to find a feasible solution (UB) for any of the three formulations within the prespecified time limit. The lower bound column (LB) provides the bound that CPLEX returns at termination. In fact, we direct CPLEX to solve the LP relaxation for each of the three formulations, but CPLEX was not able to return a solution for any of the formulations within the 10-hour time limit, regardless of the LP algorithm chosen.

Table 1: Computational statistics for the fullspace problem with 4 representative days

Formulation	Int Var	Bin Var	Cont Var	Constraints	UB (\$10 <sup>9</sup> )	LB (\$10 <sup>9</sup> )	Wall time (sec)
big-M	274,920	2,800	564,826	1,543,966	-	21.13	36,000
alternative big-M	274,920	2,800	1,102,426	2,081,566	-	21.13	36,000
hull	274,920	2,800	833,626	2,081,566	-	281.73	36,000

We also test the two decomposition algorithms described in section 4. The nested Benders decomposition is implemented in Pyomo/Python (Hart et al., 2011). The tailored Benders decomposition implementation is from CPLEX (Bonami et al., 2020), which is called using the Pyomo persistent solver interface (Siirola, 2017). The computational results of the two proposed decomposition algorithms are shown in Table 2.

The tailored Benders decomposition algorithm is able to solve all the three formulations to within 1% optimality gap within 10,000 seconds.

For the nested Benders decomposition, we observe that the forward pass with integrality constraints is expensive to solve. Therefore, we make a change

in the implementation so that we first use the nested Benders decomposition algorithm to solve the LP relaxation of the problem until the LP relaxation is solved to optimality or we reach the time limit of 10 hours or the iteration limit of 100. Then we perform one single forward pass with the integrality constraints to obtain a feasible solution. Although the nested Benders decomposition can obtain an upper bound and a lower bound to all the three formulations, the optimality gaps are large compared to the results from the tailored Benders decomposition. In fact, in none of the three formulations is the nested Benders decomposition able to solve the LP relaxation of the problems to optimality within the time limit. Note that the performance of the nested Benders decomposition is quite different from the numerical experiments on the GEP model performed by Lara et al. (2018) for the GEP model where the nested Benders decomposition performs well. The reason for this difference could be due to the complication brought by the transmission expansion constraints and the DC power flow equations of the GTEP model. As a result, the subproblems become larger and more dual degenerate, which makes nested Benders decomposition take not only more time to solve each iteration but also more iterations to converge.

Table 2: Computational results of the two proposed decomposition algorithms using different formulations

Algorithm	Formulation	UB (\$10 <sup>9</sup> )	LB (\$10 <sup>9</sup> )	Gap	Wall time (secs)
tailored Benders	big-M	283.7	282.6	0.38%	5,115
tailored Benders	alternative big-M	283.9	281.6	0.82%	3,693
tailored Benders	hull	282.6	280.6	0.71%	8,418
nested Benders	big-M	295.7	268.9	9.98%	53,682
nested Benders	alternative big-M	294.2	265.5	10.81%	43,389
nested Benders	hull	288.0	269.3	6.97%	37,577

From this numerical experiment, the tailored Benders decomposition algorithm with the alternative big-M formulation proves to be the fastest. We adopt this algorithm-formulation combination for the rest of the experiments in this paper.

Some additional computational statistics on this problem with 4 representative days including the problem sizes after presolve, the performance curves

for the nested Benders decomposition, the sizes of the Benders master problem and subproblem are shown in Appendix C of the supplementary material.

### 5.3. Results with 15 representative days

We improve the fidelity of the model by increasing the number of representative days to 15. The 15 representative day model with the alternative big-M formulation has 2,800 binary variables, 1,024,680 general integer variables, 4,120,606 continuous variables, and 7,787,266 constraints. The proposed tailored Benders decomposition algorithm is able to solve the problem in 33,207 seconds with an upper bound of 301.1 ( $\$10^9$ ), a lower bound of 299.9 ( $\$10^9$ ) and an optimality gap of 0.4%.

The capacities of different generation technologies from 2019 to 2038 are shown in Figure 3. The results include high capacities of solar and wind. The aggregated natural gas capacity of the five regions increases in the first few years, reaches its peak in 2024 and gradually decreases afterwards due to the retirement of old generators and the increase in carbon tax, which makes the natural gas generators less competitive compared with solar and wind generators. The nuclear capacities are unchanged throughout the planning horizon. The coal capacities are unchanged in the first few years and start decreasing in 2029 because of reaching their nominal lifetimes. No storage unit is installed. Therefore, the renewable generation when the net load is negative has to be curtailed. The total discounted renewable curtailment cost is 1.64 billion in 20 years.

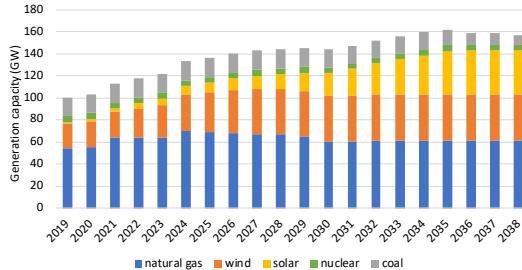


Figure 3: Aggregated generation expansion results

Geographically, most of the solar and wind capacity additions are projected to take place in the West and Panhandle regions because the capacity factors for solar and wind are higher in these two regions. The projected capacity for natural gas in the four regions, i.e., Coast, Northeast, South, and West, are shown in Figure 4. It can be seen that most natural gas expansions are expected to take place in the Northeast and Coast regions where the absolute increase in load is high and capacity factors for renewables are relatively low. In the West region, where the absolute load increase is low and the capacity factors for renewable generation are high, we observe very marginal changes in natural gas capacity. In the South region, natural gas capacity increases in the first few years and reaches a peak in 2025. After 2025, natural gas capacity decreases over the years due to the retirement of old generators. The load growth in South is satisfied by power transfers from West region, which we analyze below.

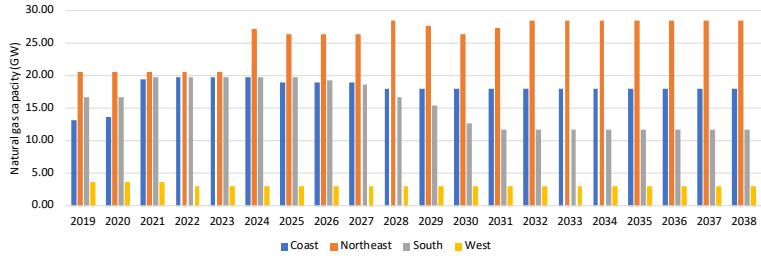


Figure 4: Projected capacities for natural gas in Coast, Northeast, South, and West

The number of transmission lines built over the planning horizon are shown in Figure 5. Most of the transmission lines are built for Northeast-Panhandle and South-West in order to transfer the power generated by the renewable sources in West and Panhandle to other regions. Note that we assume that no transmission lines are built *a priori*. It is clear that there are correlations between the geographical locations of the generation technologies and the transmission expansion decisions.

Figure 6 shows the aggregated power flow through all the installed transmission lines at a peak load time period ( $t = 20, d = 15, s = 24$ ). The directions and the magnitudes of the power flows are represented by the red arrows and

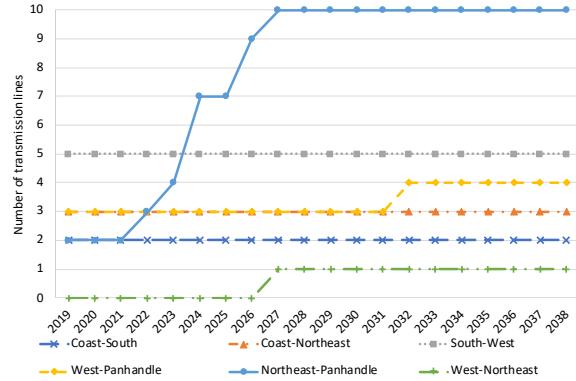


Figure 5: Transmission expansion results

numbers in this Figure. The most significant power flows are from Panhandle to Northeast and from West to South due the surplus of their renewable energy generation.

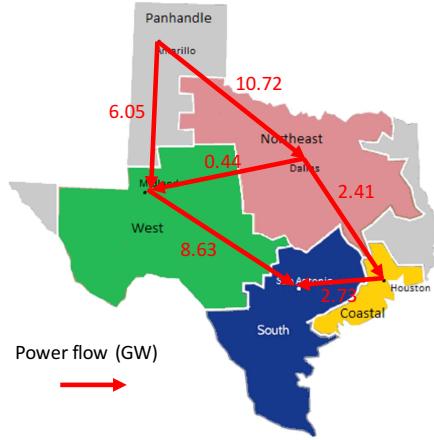


Figure 6: Aggregated power flow directions and magnitudes for all the transmission lines at  $t = 20$ ,  $d = 15$ ,  $s = 24$

#### 5.4. Sensitivity analysis of input generator data

The generation capacity expansion planning results can be sensitive to the forecast of future capital cost and operating cost of the potential generating

units. Recall that generator cost data for the previous analysis are from the National Renewable Energy Laboratory (NREL), available in the 2016 Annual Technology Baseline (ATB) Spreadsheet (Cole et al., 2016). We switch the capital cost and operating cost data for all the generators to internal data from IHS Markit and re-run the Benders decomposition experiment with 15 representative days. The proposed tailored Benders decomposition algorithm is able to solve the problem in 51,445 seconds with an upper bound of 294.4 ( $\$10^9$ ), a lower bound of 293.7 ( $\$10^9$ ) and an optimality gap of 0.2%. The increase in computational time compared with using NREL data is 67%.

The generation expansion results are shown in Figure 7. The major change is that the IHS Markit results favors more wind and less solar compared with the results shown in Figure 3. To give some insight on why this happens, we find that the overnight cost of the wind generating units in the IHS Markit dataset is 25% lower than that in the NREL dataset in the first year of the planning horizon. In contrast, the overnight cost for the PV generating units is only 11% lower in the first year.

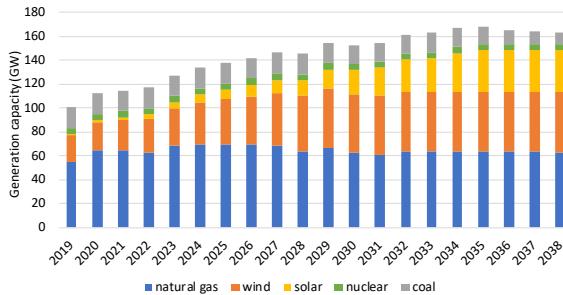


Figure 7: Generation expansion results using generating unit cost data from IHS Markit

## 6. Conclusions

In this paper, we address generation and transmission expansion planning (GTEP) problem by developing MILP formulations and solution techniques. We consider both thermal and renewable technologies as expansion candidates.

Operating and transmission constraints are included in the model, which leads to large scale problems. To limit the size of the GTEP model, several simplifications are made. We aggregate the generators that use the same technology assuming that they have the same design parameters. We also spatially aggregate regions with similar climate and load profiles. For example, the ERCOT is divided into five regions. Each region represents one bus in the power flow model. Therefore, we only consider the expansion of tielines between regions. In terms of temporal representation, we select some representative days with hourly load and capacity factor data. The representative days are selected by a clustering algorithm such as  $k$ -means clustering.

The model is a multi-scale MILP model with both investment decisions and operating decisions. We compare three different formulations for transmission expansion, i.e., the big-M formulation, the hull formulation and the alternative big-M formulation. We prove that the alternative big-M (ABM) formulation has the same feasible region as the big-M formulation (BM) when projected onto the space of the variables involved in the big-M formulation. Computational experiments are performed as well for the three formulations, but it is hard to identify a clear winner among the three formulations.

Two solution techniques, a nested Benders decomposition algorithm and a tailored Benders decomposition algorithm, are proposed. Both algorithms decompose the planning problem by year. The nested Benders decomposition solves each year sequentially in a forward and backward pass manner. The Benders decomposition defines a master problem that deals with the investment decisions and a number of subproblems corresponding to representative operating decisions for a given year. The tailored Benders decomposition algorithm outperforms the nested Benders decomposition one in our computational experiments.

An ERCOT case study is used to demonstrate the GTEP model and the solution techniques. The tailored Benders decomposition is able to solve the 20 year planning problem with 15 representative days. The capacity expansion mix for ERCOT will mainly include solar and wind capacities in the West and

Panhandle regions. The transmission lines are mainly built to transfer power from solar and wind rich regions to the South and Northeast regions of ERCOT, which shows that the generation and transmission decisions are correlated. Co-optimization of generation and transmission has the potential of bring additional value to the system operator/regulator than solving the two planning problems independently.

Several future directions can be pursued. First, the deterministic GTEP model can be extended to a multi-stage stochastic programming framework by considering uncertainties in load growth rate, carbon tax, etc. Second, the benefit of using DC power flow compared with the network flow model is limited in a small scale problem. It is worth testing the model in a problem with larger number of nodes to test the scalability of the proposed approach. Third, the computational performance of both decomposition algorithms can be improved using warm-start techniques where a number of cuts can be generated at the beginning based on the solutions obtained from some heuristics. Fourth, the reliability of the system can be improved by adding contingency constraints.

### Acknowledgements

The authors gratefully acknowledge financial support from the Center of Advanced Process Decion-making at Carnegie Mellon University and from the Department of Energy as part of the IDAES Project. Can Li would like to thank Dr. Cristiana Lara for providing some of the input data and the Python/Pyomo code for the GEP model.

Aghaei, J., Amjadi, N., Baharvandi, A., and Akbari, M.-A. (2014). Generation and transmission expansion planning: MILP-based probabilistic model. *IEEE Transactions on Power Systems*, 29(4):1592–1601.

Alguacil, N., Motto, A. L., and Conejo, A. J. (2003). Transmission expansion planning: a mixed-integer LP approach. *IEEE Transactions on Power Systems*, 18(3):1070–1077.

- Andrade, J. and Baldick, R. (2016). Estimation of transmission costs for new generation. *White Paper UTEI/2016-08-1*.
- Bahiense, L., Oliveira, G. C., Pereira, M., and Granville, S. (2001). A mixed integer disjunctive model for transmission network expansion. *IEEE Transactions on Power Systems*, 16(3):560–565.
- Baringo, L. and Baringo, A. (2017). A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. *IEEE Transactions on Power Systems*, 33(1):792–802.
- Birchfield, A. B., Xu, T., Gegner, K. M., Shetye, K. S., and Overbye, T. J. (2016). Grid structural characteristics as validation criteria for synthetic networks. *IEEE Transactions on power systems*, 32(4):3258–3265.
- Birge, J. R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33(5):989–1007.
- Blair, N., Dobos, A. P., Freeman, J., Neises, T., Wagner, M., Ferguson, T., Gilman, P., and Janzou, S. (2014). System advisor model, sam 2014.1. 14: General description. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Bonami, P., Salvagnin, D., and Tramontani, A. (2020). Implementing automatic Benders decomposition in a modern MIP solver. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 78–90. Springer.
- Cole, W., Kurup, P., Hand, M., Feldman, D., Sigrin, B., Lantz, E., Stehly, T., Augustine, C., Turchi, C., O’Connor, P., et al. (2016). 2016 annual technology baseline (atb). Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Conejo, A. J., Baringo, L., Kazempour, S. J., and Sissiqui, A. S. (2016). *Investment in Electricity Generation and Transmission - Decision Making under Uncertainty*. Springer International Publishing.

- Ding, J. and Soman, A. (2010). A long-term investment planning model for mixed energy infrastructure integrated with renewable energy. In *2010 IEEE Green Technologies Conference*, pages 1–10.
- Draxl, C., Clifton, A., Hodge, B.-M., and McCaa, J. (2015). The wind integration national dataset (wind) toolkit. *Applied Energy*, 151:355–366.
- ERCOT (2106). Hourly Load Data Archives.
- Flores-Quiroz, A., Palma-Behnke, R., Zakeri, G., and Moreno, R. (2016). A column generation approach for solving generation expansion planning problems with high renewable energy penetration. *Electric Power Systems Research*, 136:232 – 241.
- Frank, S. and Rebennack, S. (2016). An introduction to optimal power flow: Theory, formulation, and examples. *IIE Transactions*, 48(12):1172–1197.
- Gacitua, L., Gallegos, P., Henriquez-Auba, R., Lorca, A., Negrete-Pincetic, M., Olivares, D., Valenzuela, A., and Wenzel, G. (2018). A comprehensive review on expansion planning: Models and tools for energy policy analysis. *Renewable and Sustainable Energy Reviews*, 98:346–360.
- Grossmann, I. E. and Trespalacios, F. (2013). Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. *AIChe Journal*, 59(9):3276–3295.
- Hart, W. E., Watson, J.-P., and Woodruff, D. L. (2011). Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3(3):219.
- Hemmati, R., Hooshmand, R.-A., and Khodabakhshian, A. (2013). State-of-the-art of transmission expansion planning: Comprehensive review. *Renewable and Sustainable Energy Reviews*, 23:312–319.
- Koltsaklis, N. E. and Dagoumas, A. S. (2018). State-of-the-art generation expansion planning: A review. *Applied Energy*, 230:563–589.

- Koltsaklis, N. E. and Georgiadis, M. C. (2015). A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. *Applied Energy*, 158:310 – 331.
- Krishnan, V., Ho, J., Hobbs, B. F., Liu, A. L., McCalley, J. D., Shahidehpour, M., and Zheng, Q. P. (2016). Co-optimization of electricity transmission and generation resources for planning and policy analysis: review of concepts and modeling approaches. *Energy Systems*, 7(2):297–332.
- Lara, C. L., Mallapragada, D. S., Papageorgiou, D. J., Venkatesh, A., and Grossmann, I. E. (2018). Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. *European Journal of Operational Research*, 271(3):1037–1054.
- Lara, C. L., Siirola, J. D., and Grossmann, I. E. (2019). Electric power infrastructure planning under uncertainty: stochastic dual dynamic integer programming (SDDiP) and parallelization scheme. *Optimization and Engineering*, pages 1–39.
- Le Cadre, H., Papavasiliou, A., and Smeers, Y. (2015). Wind farm portfolio optimization under network capacity constraints. *European Journal of Operational Research*, 247(2):560–574.
- Liu, Y., Sioshansi, R., and Conejo, A. J. (2017). Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. *IEEE Transactions on Power Systems*, 33(1):781–791.
- Lohmann, T. and Rebennack, S. (2017). Tailored Benders decomposition for a long-term power expansion model with short-term demand response. *Management Science*, 63(6):2027–2048.
- Mallapragada, D. S., Papageorgiou, D. J., Venkatesh, A., Lara, C. L., and Grossmann, I. E. (2018). Impact of model resolution on scenario outcomes for electricity sector system expansion. *Energy*, 163:1231–1244.

- McFarland, J. R., Fawcett, A. A., Morris, A. C., Reilly, J. M., and Wilcoxen, P. J. (2018). Overview of the emf 32 study on us carbon tax scenarios. *Climate Change Economics*, 9(01):1840002.
- Mejía-Giraldo, D. and McCalley, J. D. (2013). Maximizing future flexibility in electric generation portfolios. *IEEE Transactions on Power Systems*, 29(1):279–288.
- O'Neill, R. P., Krall, E. A., Hedman, K. W., and Oren, S. S. (2013). A model and approach to the challenge posed by optimal power systems planning. *Mathematical Programming*, 140(2):239–266.
- Oree, V., Hassen, S. Z. S., and Fleming, P. J. (2017). Generation expansion planning optimisation with renewable energy integration: A review. *Renewable and Sustainable Energy Reviews*, 69:790–803.
- Palmintier, B. and Webster, M. (2011). Impact of unit commitment constraints on generation expansion planning with renewables. In *2011 IEEE power and energy society general meeting*, pages 1–7. IEEE.
- Pina, A., Silva, C. A., and Ferrão, P. (2013). High-resolution modeling framework for planning electricity systems with high penetration of renewables. *Applied Energy*, 112:215 – 223.
- Poncelet, K., Delarue, E., Duerinck, J., Six, D., and D'haeseleer, W. (2014). The importance of integrating the variability of renewables in long-term energy planning models. *KU Leuven*, pages 1–18.
- Pozo, D., Sauma, E. E., and Contreras, J. (2012). A three-level static MILP model for generation and transmission expansion planning. *IEEE Transactions on Power systems*, 28(1):202–210.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., and Rei, W. (2016). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3):801–817.

- Sadeghi, H., Rashidinejad, M., and Abdollahi, A. (2017). A comprehensive sequential review study through the generation expansion planning. *Renewable and Sustainable Energy Reviews*, 67:1369–1394.
- Schmidt, O., Hawkes, A., Gambhir, A., and Staffell, I. (2017). The future cost of electrical energy storage based on experience rates. *Nature Energy*, 2(17110):1–1.
- Scott, I. J., Carvalho, P. M., Botterud, A., and Silva, C. A. (2019). Clustering representative days for power systems generation expansion planning: Capturing the effects of variable renewables and energy storage. *Applied Energy*, 253:113603.
- Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., and Shelby, J. (2018). The national solar radiation data base (nsrdb). *Renewable and Sustainable Energy Reviews*, 89:51–60.
- Short, W., Sullivan, P., Mai, T., Mowers, M., Uriarte, C., Blair, N., Heimiller, D., and Martinez, A. (2011). Regional energy deployment system (reeds). Technical Report December, National Renewable Technology Laboratory (NREL).
- Shortt, A. and O’Malley, M. (2010). Impact of variable generation in generation resource planning models. In *IEEE PES General Meeting*, pages 1–6.
- Siirola, J. D. (2017). Pyomo: Introduction & IDAES development. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).
- Teichgraeber, H. and Brandt, A. R. (2019). Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Applied Energy*, 239:1283–1293.
- Ude, N. G., Yskandar, H., and Graham, R. C. (2019). A comprehensive state-of-the-art survey on the transmission network expansion planning optimization algorithms. *IEEE Access*, 7:123158–123181.

U.S. Energy Information Administration (2019). Annual Energy Outlook 2019.

<https://www.eia.gov/outlooks/aoe/pdf/aoe2019.pdf>.

- Zhang, H., Heydt, G. T., Vittal, V., and Quintero, J. (2013). An improved network model for transmission expansion planning considering reactive power and network losses. *IEEE Transactions on Power Systems*, 28(3):3471–3479.
- Zou, J., Ahmed, S., and Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming*, 175(1-2):461–502.

## Appendix A

We provide a succinct version of Appendix A here for easy reading. The detailed nomenclature and descriptions can be found in supplementary material.

$$\begin{aligned} \sum_i (p_{i,r,t,d,s}) + \sum_{l|r(l)=r} p_{l,t,d,s}^{\text{flow}} - \sum_{l|s(l)=r} p_{l,t,d,s}^{\text{flow}} + \sum_j p_{j,r,t,d,s}^{\text{discharge}} \\ = L_{r,t,d,s} + \sum_j p_{j,r,t,d,s}^{\text{charge}} + cur_{r,t,d,s} \quad \forall r, t, d, s \end{aligned} \quad (\text{A.1})$$

$$p_{i,r,t,d,s} = Qg_{i,r}^{\text{np}} \cdot Cf_{i,r,t,d,s} \cdot ngo_{i,r,t}^{\text{rn}} \quad \forall i \in \mathcal{I}_r^{\text{RN}}, r, t, d, s \quad (\text{A.2})$$

$$u_{i,r,t,d,s} = u_{i,r,t,d,s-1} + su_{i,r,t,d,s} - sd_{i,r,t,d,s} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \quad (\text{A.3})$$

$$\begin{aligned} p_{i,r,t,d,s} - p_{i,r,t,d,s-1} \leq & Ru_i^{\max} \cdot Hs \cdot Qg_{i,r}^{\text{np}} \cdot (u_{i,r,t,d,s} - su_{i,r,t,d,s}) \\ & + \max(Pg_i^{\min}, Ru_i^{\max} \cdot Hs) \cdot Qg_{i,r}^{\text{np}} \cdot su_{i,r,t,d,s} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} p_{i,r,t,d,s-1} - p_{i,r,t,d,s} \leq & Rd_i^{\max} \cdot Hs \cdot Qg_{i,r}^{\text{np}} \cdot (u_{i,r,t,d,s} - su_{i,r,t,d,s}) \\ & + \max(Pg_i^{\min}, Rd_i^{\max} \cdot Hs) \cdot Qg_{i,r}^{\text{np}} \cdot sd_{i,r,t,d,s} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \end{aligned} \quad (\text{A.5})$$

$$u_{i,r,t,d,s} \cdot Pg_i^{\min} \cdot Qg_{i,r}^{\text{np}} \leq p_{i,r,t,d,s} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \quad (\text{A.6})$$

$$p_{i,r,t,d,s} + q_{i,r,t,d,s}^{\text{spin}} \leq u_{i,r,t,d,s} \cdot Qg_{i,r}^{\text{np}} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \quad (\text{A.7})$$

$$\sum_{i \in \mathcal{I}_r^{\text{TH}}} (q_{i,r,t,d,s}^{\text{spin}} + q_{i,r,t,d,s}^{\text{Qstart}}) \geq Op^{\min} \cdot L_{r,t,d,s} \quad \forall r, t, d, s \quad (\text{A.8})$$

$$\sum_{i \in \mathcal{I}_r^{\text{TH}}} q_{i,r,t,d,s}^{\text{spin}} \geq Spin^{\min} \cdot L_{r,t,d,s} \quad \forall r, t, d, s \quad (\text{A.9})$$

$$q_{i,r,t,d,s}^{\text{spin}} \leq u_{i,r,t,d,s} \cdot Qg_{i,r}^{\text{np}} \cdot Frac_i^{\text{spin}} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \quad (\text{A.10})$$

$$q_{i,r,t,d,s}^{\text{Qstart}} \leq (ngo_{i,r,t}^{\text{th}} - u_{i,r,t,d,s}) \cdot Qg_{i,r}^{\text{np}} \cdot Frac_i^{\text{Qstart}} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t, d, s \quad (\text{A.11})$$

$$\sum_{i \in \mathcal{I}_r^{\text{RN}}} \sum_r (Qg_{i,r}^{\text{np}} \cdot Q_i^{\text{v}} \cdot ngo_{i,r,t}^{\text{rn}}) + \sum_{i \in \mathcal{I}_r^{\text{TH}}} \sum_r (Qg_{i,r}^{\text{np}} \cdot ngo_{i,r,t}^{\text{th}}) \geq (1 + R_t^{\min}) \cdot L_t^{\max} \quad \forall t \quad (\text{A.12})$$

$$\sum_d \sum_s \left[ W_d \cdot Hs \cdot \left( \sum_r \left( \sum_{i \in \mathcal{I}_r^{\text{RN}}} p_{i,r,t,d,s} - cu_{r,t,d,s} \right) \right) \right] + def_t^{\text{rn}} \geq RN_t^{\min} \cdot ED_t \quad \forall t \quad (\text{A.13})$$

$$ED_t = \sum_r \sum_d \sum_s (W_d \cdot Hs \cdot L_{r,t,d,s})$$

$$\sum_r ngb_{i,r,t}^{\text{rn}} \leq Q_{i,t}^{\text{inst},\text{UB}} / Qg_{i,r}^{\text{np}} \quad \forall i \in \mathcal{I}_r^{\text{Rnew}}, t \quad (\text{A.14})$$

$$\sum_r ngb_{i,r,t}^{\text{th}} \leq Q_{i,t}^{\text{inst},\text{UB}} / Qg_{i,r}^{\text{np}} \quad \forall i \in \mathcal{I}_r^{\text{Tnew}}, t \quad (\text{A.15})$$

$$ngo_{i,r,t}^{\text{rn}} = Ng_{i,r}^{\text{Rold}} + ngb_{i,r,t}^{\text{rn}} - ngr_{i,r,t}^{\text{rn}} \quad \forall i \in \mathcal{I}_r^{\text{RN}}, r, t = 1 \quad (\text{A.16})$$

$$ngo_{i,r,t}^{\text{rn}} = ngo_{i,r,t-1}^{\text{rn}} + ngb_{i,r,t}^{\text{rn}} - ngr_{i,r,t}^{\text{rn}} \quad \forall i \in \mathcal{I}_r^{\text{RN}}, r, t > 1 \quad (\text{A.17})$$

$$Ng_{i,r,t}^{\text{r}} = ngr_{i,r,t}^{\text{rn}} + nge_{i,r,t}^{\text{rn}} \quad \forall i \in \mathcal{I}_r^{\text{Rold}}, r, t \quad (\text{A.18})$$

$$ngo_{i,r,t}^{\text{th}} = Ng_{i,r}^{\text{Told}} + ngb_{i,r,t}^{\text{th}} - ngr_{i,r,t}^{\text{th}} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t = 1 \quad (\text{A.19})$$

$$ngo_{i,r,t}^{\text{th}} = ngo_{i,r,t-1}^{\text{th}} + ngb_{i,r,t}^{\text{th}} - ngr_{i,r,t}^{\text{th}} \quad \forall i \in \mathcal{I}_r^{\text{TH}}, r, t > 1 \quad (\text{A.20})$$

$$Ng_{i,r,t}^{\text{r}} = ngr_{i,r,t}^{\text{th}} + nge_{i,r,t}^{\text{th}} \quad \forall i \in \mathcal{I}_r^{\text{Told}}, r, t \quad (\text{A.21})$$

$$u_{i,r,t,d,s} \leq ngo_{i,r,t}^{\text{th}} \quad \forall i \in \mathcal{I}_r^{\text{Tnew}}, r, t, d, s \quad (\text{A.22})$$

$$nso_{j,r,t} = Ns_{j,r} + nsb_{j,r,t} - nsr_{s,r,t} \quad \forall j, r, t = 1 \quad (\text{A.23})$$

$$nso_{j,r,t} = nso_{j,r,t-1} + nsb_{j,r,t} - nsr_{j,r,t} \quad \forall j, r, t > 1 \quad (\text{A.24})$$

$$Charge_j^{\min} \cdot nso_{j,r,t} \leq p_{j,r,t,d,s}^{\text{charge}} \leq Charge_j^{\max} \cdot nso_{j,r,t} \quad \forall j, r, t, d, s \quad (\text{A.25})$$

$$Discharge_j^{\min} \cdot nso_{j,r,t} \leq p_{j,r,t,d,s}^{\text{discharge}} \leq Discharge_j^{\max} \cdot nso_{j,r,t} \quad \forall j, r, t, d, s \quad (\text{A.26})$$

$$Storage_j^{\min} \cdot nso_{j,r,t} \leq p_{j,r,t,d,s}^{\text{level}} \leq Storage_j^{\max} \cdot nso_{j,r,t} \quad \forall j, r, t, d, s \quad (\text{A.27})$$

$$p_{j,r,t,d,s}^{\text{level}} = p_{j,r,t,d,s-1}^{\text{level}} + \eta_j^{\text{charge}} \cdot p_{j,r,t,d,s}^{\text{charge}} + p_{j,r,t,d,s}^{\text{discharge}} / \eta_j^{\text{discharge}} \quad \forall j, r, t, d, s > 1 \quad (\text{A.28})$$

$$p_{j,r,t,d,s}^{\text{level}} = p_{j,r,t,d}^{\text{level},0} + \eta_j^{\text{charge}} \cdot p_{j,r,t,d,s}^{\text{charge}} + p_{j,r,t,d,s}^{\text{discharge}} / \eta_j^{\text{discharge}} \quad \forall j, r, t, d, s = 1 \quad (\text{A.29})$$

$$p_{j,r,t,d}^{\text{level},0} = 0.5 \cdot Storage_j^{\max} \cdot nso_{j,r,t} \quad \forall j, r, t, d \quad (\text{A.30})$$

$$p_{j,r,t,d,s}^{\text{level}} = 0.5 \cdot Storage_j^{\max} \cdot nso_{j,r,t} \quad \forall j, r, t, d, s = S \quad (\text{A.31})$$

$$\min_t \Phi = \sum_t (\Phi_t^{\text{opex}} + \Phi_t^{\text{capex}} + \Phi_t^{\text{PEN}}) \quad (\text{A.32})$$