Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Joint Planning of Energy Storage and Transmission for Wind Energy Generation

Wei Qi, Yong Liang, Zuo-Jun Max Shen

Operations Research
Articles in Advance, pp. 1–14
ISSN 0030-364X (print) | ISSN 1526-5463 (online)

http://dx.doi.org/10.1287/opre.2015.1444

Joint Planning of Energy Storage and Transmission for Wind Energy Generation

Wei Qi
Department of Industrial Engineering, Tsinghua University, Beijing 100084, China, qiwei.0216@gmail.com

Yong Liang
School of Economics and Management, Tsinghua University, Beijing 100084, China, liangyong@sem.tsinghua.edu.cn

Zuo-Jun Max Shen
Department of Industrial Engineering and Operations Research and Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, maxshen@berkeley.edu

Regions with abundant wind resources usually have no ready access to the existing electric grid. However, building transmission lines that instantaneously deliver all geographically distributed wind energy can be costly. Energy storage (ES) systems can help reduce the cost of bridging wind farms and grids and mitigate the intermittency of wind outputs. In this paper, we propose models of transmission network planning with colocation of ES systems. Our models determine the sizes and sites of ES systems as well as the associated topology and capacity of the transmission network under the feed-in-tariff policy instrument. We first formulate a location model as a mixed-integer second-order-conic program to solve for the ES-transmission network design with uncapacitated storage. Then we propose a method to choose ES sizes by deriving a closed-form upper bound. The major insight is that, in most cases, using even small-sized ES systems can significantly reduce the total expected cost, but their marginal values diminish faster than those of the transmission lines as their capacities expand. Despite uncertainties in climate, technologies, and construction costs, the cost-efficient infrastructure layout is remarkably robust. We also identify the major bottleneck cost factors for different forms of ES technologies.

Keywords: facility location; wind energy; energy storage; infrastructure planning.
Subject classifications: Environment; facilities/equipment planning: location; natural resources: energy.
Area of review: Environment, Energy, and Sustainability.
History: Received July 2014; revisions received May 2015, August 2015; accepted September 2015. Published online in *Articles in Advance*.

1. Introduction

Renewable energy, such as wind energy, is the key to a sustainable energy future. Drivers for the renewables include alleviated dependence on fossil-fuel power and nuclear power, reduced environmental hazards, and prospective cheaper energy production (EIA 2014). Governments around the world have widely released targets to push the adoption of renewable energy. For example, a collaborative effort has been made to explore a scenario in which wind provides 20% of U.S. electricity by 2030 (DOE 2008); China plans its non-fuel energy to account for 15% of its total consumption by 2020 and has led the expansion of wind power capacity (REN21 2011). In the mean time, leading companies in the information technology (IT) sector, such as Apple, Google, and Facebook, are taking significant steps to power their data centers with an increasing percent of renewable energy (Greenpeace 2014).

Nonetheless, wind energy infrastructure planners are facing major challenges when they are trying to meet these ambitious goals. Firstly, wind resources are geographically distributed. In the initial phase, when multiple wind farms are approved to be built, the planners need to carefully design the transmission network that is usually more complex than one single line. Secondly, most of the high-quality wind resources in North America and Asia are not near major load centers and cannot be directly integrated into the existing transmission network (Denholm and Sioshansi 2009; Elliott et al. 1991, 2001; Greenblatt et al. 2007). As a result, dedicated long-distance transmission lines have to be built to deliver electricity from remote wind farms. Thirdly, the intermittent nature of wind necessitates high-capacity but lightly-loaded transmission lines, which otherwise would result in significant generation curtailment. For example, Southern California Edison reported curtailed wind energy generation of about 15 MW for 6%–8% of the time as of 2010 because of transmission constraints (Rogers et al. 2010).

We try to address these challenges by proposing models of transmission network planning with colocation of energy storage (ES) systems. The primary function of an ES system is to decrease the variability of wind energy generation by absorbing/discharging electricity when wind power output mismatches the rated transmission capacity or power demand. The extensive value of colocating ES with wind energy generation has been reported in EAC (2008) and EPRI-DOE (2003). In this paper, our goal is to develop models and solution methods to determine sizes and sites of ES systems as well as the associated topology and capacity of the transmission network. As a result, wind energy from these geographically distributed wind farms can be effectively tapped with minimum infrastructure investment cost and energy loss. In doing so, we also try to understand how to best exploit the value of using ES for future renewable energy production.

Our model and analysis are based on the following problem settings: a set of sites in a region with abundant wind resources have been selected as wind farms. These sites are located in desolated areas that have no ready access to main transmission infrastructure (see Talinli et al. 2011 for practical considerations in siting wind farms). A planner of the local government or a utility company is to design a network of energy storage systems and transmission lines (hereafter referred to as ES-transmission network) to connect the wind farms to a single load center (e.g., a town) or a substation of the region; or, an IT company aims to power its 120 MW data center with 100 % wind energy. We consider the network topology to be radial. That is, wind outputs from different farms are first transmitted to junction sites with or without ES, and then the pooled power at each junction site flows to the load center. The radial transmission network is widely adopted in practice to tap remote wind resources, such as in Southern California and Atlantic offshore zones (California ISO 2013 and AWC 2014).

In addition, we assume that the region implements feed-in-tariffs (FIT) policy, which guarantees a long-term contract for renewable power producers to sell their electricity at a fixed price (Mendonça et al. 2009). As a result, wind farms have no price arbitrage incentive and it is optimal to deliver as much energy as transmission capacities permit. Among the existing policy mechanisms, FIT is particularly effective to foster initial adoption of renewable energy and fits well with the practice in most of the world's major electricity markets, where governments enforce the purchase price of electricity to be higher than its energy production cost or subsidize wind energy generation utilities to attract them to enter the market (Alizamir et al. 2012 and REN21 2011).

Our first model considers the case where ES systems are assumed to have sufficient energy capacity to accommodate intermittent surplus wind output. We first derive two optimal transmission line capacities as functions of wind characteristics for a single wind farm with and without ES being coupled, respectively. Then we use these optimal quantities to formulate a model to design an ES-transmission network. The model is in the form of a mixed-integer second-order-conic program (MISOCP), which can be efficiently solved by commercial software. Our second model considers the sizing problem of ES. We derive a closed-form upper bound of the expected energy overflow due to the capacity limit, as a function of ES and transmission capacities. In the above models, following one similar assumption of Kim and Powell (2011), we approximate the hourly and daily wind output by uniform distribution. The numerical experiments suggest that the approximation error is small. Combining these models, the infrastructure planner obtains both lower and upper bounds of the expected minimum capital and operational cost of the network. The gap between the two bounds is reasonably small.

The contributions of our paper are as follows: (1) To the best of our knowledge, this paper is the first attempt to provide infrastructure planners with models and solution approaches to jointly plan the sites and the sizes of ES systems and transmission lines for distributed wind resources. (2) We develop quantitative models and managerial insights to help planners understand the value and cost of using ES. We analyze the dual effects of using ES, that is, saving transmission capacity by reducing output variability versus incurring energy loss due to friction and overflow (as an in-depth quantitative extension to the discussion in Denholm and Sioshansi 2009). We find that, in most cases, using even small-sized ES systems can significantly reduce the total expected cost, but their marginal values diminish faster than those of the transmission lines as their capacities expand. We also identify the bottleneck cost factors for different forms of ES technologies. For example, for compressed air storage systems, it is more beneficial to improve their energy conversion efficiency than to reduce their per-unit capacity cost. These insights can be used to make long-term investment decisions as technology advancements bring down ES cost. (3) Another finding is that the layout of the ES-transmission network that we obtain is robust against uncertainties such as FIT rates adjustments, technology advancements, climate changes, and construction material cost fluctuations. Hence, planners can determine the infrastructure layout long before these uncertainties are resolved, without worrying about costly reconfiguration of the network. (4) We also incorporate major wind characteristics into infrastructure planning. In particular, our models capture the nature of wind energy such as hourly and daily intermittence, spatial correlation, and the variability pooling effect, which are all important factors but have not been well considered in the literature.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces basic settings of our models. Section 4 presents the infrastructure planning model with uncapacitated ES systems. Section 5 derives an upper bound of the size of an ES system and incorporates it into the planning procedure. Section 6 demonstrates our computational results and presents managerial insights into technology impact and layout robustness. Finally, §7 concludes the paper. In addition, numerical and theoretical analysis of model inaccuracy, additional structural properties, proposition proofs, and numerical experiment settings are available in the e-companion (available as supplemental material at http://dx.doi.org/10.1287/opre.2015.1444) of this paper.

2. Literature Review

There has only been a very limited number of studies that are related to the important problem of deploying ES systems for wind power delivery in the literature of transmission expansion planning (see Latorre et al. 2003, Hemmati et al. 2013 for comprehensive reviews; Taylor and Hover 2013 for recent progress on conic approximations to alternative current (AC) transmission system planning; and Moeini-Aghtaie et al. 2012, Baringo and Conejo 2012 for some recent studies that incorporate wind resources). The closest to ours are Oh (2011) and Zhang et al. (2013), both of which formulate deterministic mixed-integer linear programs to plan ES systems in an existing power grid, without wind resources being considered. In addition, Denholm and Sioshansi (2009) study how ES saves transmission cost when it is located close to wind farms. However, none of these studies address issues such as determining ES/transmission capacities while capturing wind characteristics, which are the contributions of our paper from a supply chain design perspective.

Our infrastructure planning model for wind energy generation is reminiscent of the location-inventory model proposed by Shen et al. (2003), though they are different on some fundamental aspects. Compared to warehouses in a distribution network, the ES systems incur fixed upfront cost, variable capacity cost, and nonlinear cost because of charge/discharge friction loss and overflow loss of energy. Meanwhile, whereas the retailers to be assigned to the warehouses face random custom demands, the wind farms in our problem setting are to be assigned to ES-coupled or ES-free junction sites and face intermittent wind energy outputs. For the solution approach, we formulate our planning problem in the form of a computationally tractable second-order conic program. For more applications of conic formulations, please refer to Atamtürk et al. (2012) for solving the location-inventory problem and its various extensions, Mak et al. (2013) for planning battery-swapping stations of electric vehicles, and Natarajan et al. (2010) for portfolio optimization.

3. Model Settings

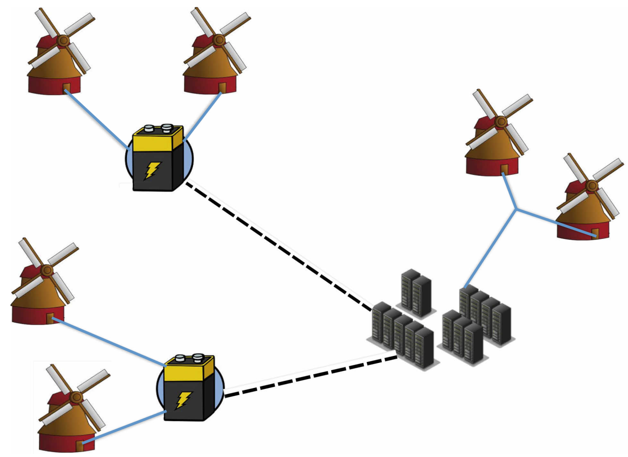
We consider a radial ES-transmission network as illustrated in Figure 1. A given set of geographically distributed wind farms are approved to be built. Each of these farms is to generate and deliver electrical power to a junction site. Then the power pooled at each junction site flows to a given common load center (or substation). Our objective is to jointly determine (a) the assignment of the wind farms to the junction sites, (b) whether to install an ES system at a selected junction site, (c) energy capacities of the ES systems, and (d) power capacities of the transmission lines.

We are interested in how the total expected cost relates to ES and transmission capacities as well as wind intermittence. Specifically, this cost breaks down to two parts:

(1) The *building cost* of ES (transmission), which consists of a fixed installation cost and a variable cost proportional to the ES (transmission) capacity. The fixed and the variable cost components of transmission are assumed to be both proportional to the length of the line.

(2) And the *energy loss*, which is incurred as (i) *friction loss*, owing to nonperfect round-trip conversion efficiency

Figure 1. (Color online) A radial ES-transmission network with economic (dashed) and ES-free (solid) lines.



when electricity is charged into and discharged from an ES system; (ii) *overflow loss*, when both an ES system and its downstream transmission line hit their maximum capacities and the surplus wind energy can be neither stored nor transmitted; or (iii) *curtailment loss*, when ES is absent and instantaneous wind power output that exceeds the downstream transmission line capacity has to be abandoned.

For the choices of transmission capacity, we consider two types of transmission lines. First, between ES-equipped junction sites and the load center, we choose *economic lines*. An economic line and its associated ES system are complementary in reducing energy loss and their capacities need to be jointly optimized. Second, *ES-free lines* are built between the wind farms and the junction sites and from ES-free junction sites to the load center. A properly sized ES-free line strikes the balance between saving transmission building cost and reducing curtailment loss. In both cases, given the FIT instrument, wind farms have no price arbitrage incentive. It is optimal to deliver as much generated and stored energy as transmission line capacities permit.

4. Model with Uncapacitated Storage

In this section, we first derive an optimal transmission line capacity for a single wind farm colocated with an uncapacitated ES system. Then we derive an optimal transmission capacity of an ES-free line. These results then lead to an ES-transmission planning problem formulation for multiple distributed wind farms with uncapacitated ES systems. In §5 we will show that this uncapacitated case provides a reasonable approximation to the case with capacity limits. Table 1 summarizes the notation for §§4.1, 4.2, and 4.3. Parameters and functions are denoted by lowercase letters, random variables by bold lowercase letters, matrices by uppercase Greek letters, and decision variables by uppercase English letters.

Table 1. Summary of notation.

	Systems parameters
α, β	Charge and discharge efficiency of ES systems, respectively.
δ	Length (in years) of each time interval. We assume $\delta = 1$ hour $= 1/(24 \times 365)$ year hereafter.
$\mathbf{w}_t, w_t$	Random variable and its realization of the energy that a wind farm captures during interval $[t-1, t)$, respectively, where $t \in \mathbb{N}$ is the index of the intervals of length δ.
$\mathbf{l}_t$	Loss of energy during $[t-1, t)$.
$f_w(\cdot)$	Probability density function of $\mathbf{w}_t$.
μ, ϵ	Mean and interval length of the approximated uniform distribution of $\mathbf{w}_t$, respectively.
l	Length of a transmission line.
	Price and costs
p	Fixed contracted electricity selling price.
r	Annualized per-kWh building cost of ES capacity.
a	Annualized building cost of a transmission line per kW per mile.
$q = al$	Annualized building cost of a transmission line per kW.
θ, η	Dimensionless capacity cost indices of an economic line and an ES-free line, respectively.
	Decision variable
C	Maximum electrical energy that can be transmitted over a period of δ by a transmission line; C is also in the unit of power (kW) when $\delta = 1$ hour $= 1/(24 \times 365)$ year.

4.1. A Single Wind Farm with ES

Consider a basic scenario: a single wind farm is coupled with an ES system and delivers electricity through a capacitated transmission line. In this case, it is optimal to colocate the ES system with the wind farm to avoid the cost of building an ES-free transmission line between them.

Intuitively, as the transmission line capacity increases and/or the installed ES capacity increases, the building cost increases while the energy loss decreases. To quantify this trade-off, we first assume that $r = 0$ and the ES system is large enough to incur no overflow loss almost surely. The energy loss $\mathbf{l}_t$ during $[t-1, t)$ thus consists only of the friction loss:

$$\mathbf{l}_t = \begin{cases} (\mathbf{w}_t - C)(1-\alpha\beta), & \text{if } \mathbf{w}_t - C > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

At times when the wind output power exceeds the transmission line capacity, the surplus energy $(\mathbf{w}_t - C)$ is charged into and at some future time discharged from the ES system, incurring a friction loss of $(\mathbf{w}_t - C)(1-\alpha\beta)$. Otherwise, all the generated wind energy can be directly delivered. We follow the approach in Kim and Powell (2011) to assume that $\mathbf{w}_t$ is uniformly distributed: $\mathbf{w}_t \sim \text{Unif}(\mu - \epsilon/2, \mu + \epsilon/2)$. We obtain the mean μ and the interval length ϵ by matching the mean and the variance of the real wind outputs. Two reasons lead to our choice of uniform distributions over others (such as normal distributions) to approximate wind outputs. First, the uniform distribution is mathematically tractable, enabling us to derive closed-form results that are key not only to the efficient planning problem formulation, but also to the managerial insights into ES value and model suboptimality. Second, the uniform distribution, with its bounded support, is effective in characterizing wind curtailment, which results from wind turbine operations and capacitated transmission lines. This approximation is further justified by the numerical experiments in §4.3 and in §EC.1.1 in the e-companion. Note that the wind output process $\{\mathbf{w}_t\}$ can be autocorrelated and nonstationary. When $C \geqslant \mu - \epsilon/2$, the expected energy loss in $[t-1, t)$ is given by

$$\begin{aligned} \mathbb{E}[\mathbf{l}_t] &= \int_C^{\mu+\epsilon/2} (w_t - C)(1-\alpha\beta)\frac{1}{\epsilon}\, dw_t \\ &= \frac{1-\alpha\beta}{2\epsilon}\left(\mu + \frac{\epsilon}{2} - C\right)^2. \end{aligned} \tag{2}$$

Although in practice transmission capacity C can only be chosen from a finite set of discrete values, we assume C is continuous valued for model tractability, because the discrete set of candidate capacities is considerably flexible, with various line specifications available. In addition, we assume that the variable transmission capital cost is linear in C. This linear approximation is present and justified in early literature of transmission expansion planning (e.g., Kaltenbach et al. 1970, Kim et al. 1988). Recent empirical evaluation (Mason et al. 2012a, b) also suggests that transmission capital cost exhibits a significant linear relation with transmission capacity in a wide range. With these two assumptions, the expected annual variable cost due to friction loss and capital investment can be expressed as a quadratic function of C:

$$\begin{aligned} v_1(C) &= \frac{p\,\mathbb{E}[\mathbf{l}_t]}{\delta} + qC \\ &= p\frac{1-\alpha\beta}{2\epsilon\delta}\left(\mu + \frac{\epsilon}{2} - C\right)^2 + qC. \end{aligned} \tag{3}$$

It can be verified that $C = (\mu + \epsilon/2) - (\epsilon\delta q)/(p(1-\alpha\beta))$ minimizes (3). We make an additional assumption that transmission line capacity should be greater than or equal to the average wind output power; otherwise there is no steady state distribution of the storage level. Hence, the economic transmission capacity that minimizes the expected annual variable cost v_1 is given by

$$\begin{aligned} C^* &= \arg\min_{C >= \mu} v_1(C) \\ &= \max\left\{\mu, \left(\mu + \frac{\epsilon}{2}\right) - \frac{\epsilon\delta q}{p(1-\alpha\beta)}\right\} \\ &= \begin{cases} \mu + \left(\frac{1}{2} - \theta\right)\epsilon, & \text{if } \theta < \frac{1}{2}; \\ \mu, & \text{otherwise,} \end{cases} \\ \text{where } \theta &= \frac{\delta q}{p(1-\alpha\beta)}. \end{aligned} \tag{4}$$

The dimensionless number θ captures the cost associated with building transmission capacity. For example, if q is large because of long transmission distance or high unit capacity cost, or if the ES conversion is very efficient such that $\alpha\beta$ is close to 1, then θ tends to be large, indicating that building extra transmission capacity is cost-ineffective. When $\theta \geqslant \frac{1}{2}$, it is favorable to construct a line that transmits at most average wind power. It is important to notice that θ is independent from the wind characteristics. When planning the entire ES-transmission network, this invariance helps predetermine which segment of the following nonsmooth cost expression (5) to use in formulating the network design problem, before we know the assignment of the wind farms to the junction sites. Substituting (4) into (3) yields the optimal annual variable cost:

$$v_1^*(C^*) = \begin{cases} q\mu + q\left(\frac{1}{2} - \frac{1}{2}\theta\right)\epsilon, & \text{if } \theta < \frac{1}{2}; \\ q\mu + \dfrac{p(1-\alpha\beta)}{8\delta}\epsilon, & \text{otherwise.} \end{cases} \tag{5}$$

4.2. A Single Wind Farm without ES

We next look into another basic scenario where a single wind farm is connected with a transmission line without ES being colocated. In this case, the optimal transmission capacity can be directly expressed as the optimal quantile in the classic newsvendor model with stock-out cost $(p - q\delta)$ and inventory holding cost $q\delta$. An explicit derivation resembles the steps in §4.1. Specifically, the curtailment loss during $[t-1, t)$ is given by

$$\mathbf{l}_t = \begin{cases} (\mathbf{w}_t - C), & \text{if } \mathbf{w}_t - C > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Applying uniform distribution approximation again, when $C \geqslant \mu - \epsilon/2$, the expected curtailment loss is given by

$$\begin{cases} \mathbb{E}[\mathbf{l}_t] = \displaystyle\int_C^{\mu+\epsilon/2} (w_t - C)\frac{1}{\epsilon}\, dw_t \\ \qquad\quad = \dfrac{1}{2\epsilon}\left(\mu + \dfrac{\epsilon}{2} - C\right)^2. \end{cases} \tag{7}$$

The expected annual variable cost due to curtailment loss and transmission capacity investment as a function of C becomes

$$\begin{aligned} v_2(C) &= \frac{p\,\mathbb{E}[\mathbf{l}_t]}{\delta} + qC \\ &= \frac{p}{2\epsilon\delta}\left(\mu + \frac{\epsilon}{2} - C\right)^2 + qC \\ &= \frac{p}{2\epsilon\delta}C^2 + \left(-\frac{p}{\epsilon\delta}\left(\mu + \frac{\epsilon}{2}\right) + q\right)C \\ &\quad + \frac{p}{2\epsilon\delta}\left(\mu + \frac{\epsilon}{2}\right)^2. \end{aligned} \tag{8}$$

The cost-minimizing ES-free line capacity is thus given by

$$\begin{aligned} C^* &= \begin{cases} \arg\min v_2(C), & \text{if } \eta < 1; \\ 0, & \text{otherwise;} \end{cases} \\ &= \begin{cases} \mu + \left(\frac{1}{2} - \eta\right)\epsilon, & \text{if } \eta < 1; \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

$$\text{where } \eta = \delta q/p. \tag{9}$$

Similar to θ for an economic line, η is the dimensionless capacity cost index for an ES-free line, independent from wind characteristics. Larger η indicates higher levels of line capacity restriction. When $\eta \geqslant 1$, building transmission capacity is no longer profitable even when the line is fully loaded all the time, so we opt not to build the line and forgo all the wind energy. The associated optimal annual cost is

$$v_2^*(C^*) = \begin{cases} q\mu + q\left(\frac{1}{2} - \frac{1}{2}\eta\right)\epsilon, & \text{if } \theta < 1; \\ p\mu/\delta, & \text{otherwise.} \end{cases} \tag{10}$$

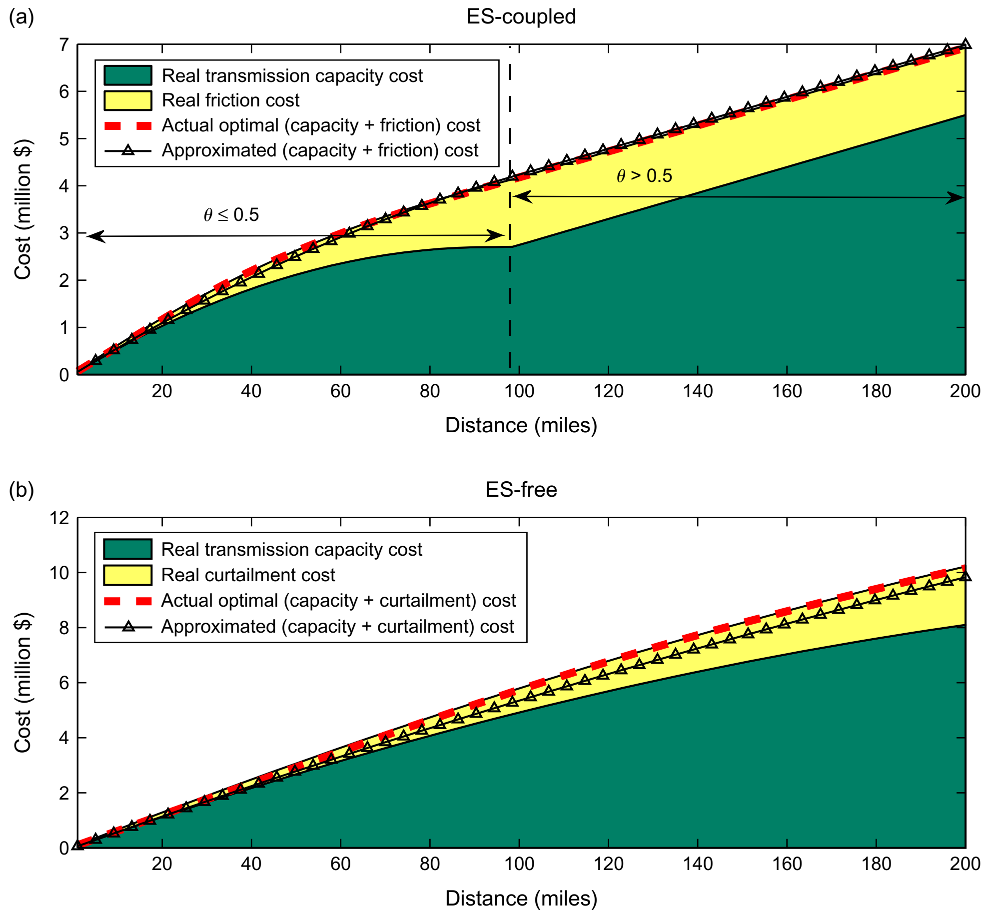
4.3. A Numerical Example

The following simple numerical example illustrates how the variable cost and the approximation error of wind output vary with respect to the transmission distance. Consider distance l ranging from 0–200 miles and set $p = \$0.08/\text{kWh}$ (which is projected to be the levelized electricity cost for new wind plants in 2020, estimated by EIA 2014), $\alpha\beta = 0.72$, and $a = \$1/\text{kW-mile}$.

Figure 2(a) and 2(b) are the area plots of the total variable cost with and without ES being colocated, respectively. We use wind output data of a modeled site from EPRI-DOE (2008) (The site ID is 24,648). In Figure 2(a), the total variable cost generated using these data with the transmission capacity C^* prescribed by (4) is shown as the sum of the friction cost and the transmission capacity cost. When $l < 98$ miles, the transmission capacity cost increases in transmission distance l yet with a decreasing rate of change, as C^* decreases to partly offset the increased capacity cost, which in turn incurs more charge/discharge friction. When $l \geqslant 98$ miles, $C^* = \mu$, and hence the friction cost reaches standstill and the transmission capacity cost increases linearly in l. In Figure 2(b), the total variable cost with C^* prescribed by (9) is shown as the sum of the curtailment cost and the transmission capacity cost. Similarly, longer transmission distance results in smaller transmission capacity and thus higher curtailment cost.

Figure 2 also suggests that the error of approximating the wind output using uniform distribution is reasonably small. We first use line search on the same data set to find the actual optimal transmission capacity and the corresponding cost. This cost, as represented by the dashed line in each plot, is close to the cost with transmission capacity C^*. In the ES-coupled (ES-free) case, the average relative gap is 3.01% (7.02%) for $l \leqslant 50$ miles and 0.22% (1.09%) for

Figure 2. (Color online) Variable costs of a single farm (a) with and (b) without ES.



$l > 50$ miles. Then we use simulated data from a uniform distribution that matches the first and the second moments of the raw data. Again, the approximated total variable cost (the line with triangles) is close to the real cost, with about a 2.5% (5%) relative gap in the ES-coupled (ES-free) case. It can be verified that the bias of the cost approximation is eliminated when $C = \mu$ regardless of the real distribution of the wind output. These facts enable us to incorporate the cost terms (5) and (10) into the planning model for multiple wind farms, which we elaborate in §4.4.

4.4. Multiple Farms

We next develop the planning model for multiple wind farms. Obviously, building ES systems at all the sites of the wind farms can be cost-inefficient. We instead try to economically select power junction sites (which can also be some wind farms) to aggregate wind outputs with or without ES systems being colocated.

Additional notation used in this subsection is summarized in Table 2. Wherever it is necessary, we add subscripts to symbols to indicate location. For instance, q_{ij} and q_j refer to capacity per unit costs of building transmission lines from wind farm i to junction site j, and from junction site j to the load center, respectively.

We assume that ES-free lines are built between wind farms and junction sites and between ES-free junction sites and the load center, with line capacities given by (9). We follow the same logic as in §§4.1 and 4.2 to use a uniform distribution to approximate the probability distribution of the curtailed wind power $\mathbf{w}_{t,ij}$, which is from farm i and faced by site j. When $\eta_{ij} < 1$, this uniform distribution has mean and interval length expressed as follows (the derivation is available in §EC.2.1 in the e-companion):

$$\mu_{ij} = \mu_i - \tfrac{1}{2}\epsilon_i \eta_{ij}^2, \tag{11a}$$

$$\epsilon_{ij} = \sqrt{(1-\eta_{ij})^3(1+3\eta_{ij})}\,\epsilon_i. \tag{11b}$$

Spatial correlation of wind speed and power has been extensively reported and used in wind forecast (Alexiadis et al. 1999); therefore, it should be explicitly modeled. Each selected junction site j faces pooled and correlated wind outputs from a subset of wind farms. Again, we apply uniform distribution approximation to this pooled wind output, i.e., $\mathbf{w}_{t,j} \sim \text{Unif}(\mu_j - \epsilon_j/2, \mu_j + \epsilon_j/2)$. Let Σ_j be the covariance matrix of wind outputs from the wind farms in I and aggregated at junction site j. Each entry $\Sigma_{ikj} = \epsilon_{ij}\rho_{ikj}\epsilon_{kj}$, where ρ_{ikj} is the correlation coefficient of the curtailed wind

Table 2. Summary of additional notation.

	Sets
I	Set of wind farms, indexed by $i \in I$.
J	Set of candidate junction sites pooling wind outputs, indexed by $j \in J$.
	Systems parameters
Σ_j	Covariance matrix of wind outputs from the wind farms in I, aggregated at site j.
ρ_{ikj}	Correlation coefficient of the wind outputs at site j from $i, k \in I$, aggregated at site j.
	Costs
h_j	Annualized fixed upfront cost to build an ES system at junction site $j \in J$.
g_{ij}, g_j	Annualized fixed construction cost of the transmission line from wind farm $i \in I$ to site $j \in J$ and from site j to the load center (or a substation), respectively.
	Decision variables
X_j	One if an ES system is built on site $j \in J$, zero otherwise.
V_j	One if site $j \in J$ is selected as a junction site with no ES system, zero otherwise.
Y_{ij}	One if wind farm $i \in I$ is assigned to junction site j with an ES system, zero otherwise.
Y_j	Vector $(Y_{1j}, Y_{2j}, \ldots, Y_{\lvert I \rvert j})^T$.
Z_{ij}	One if wind farm $i \in I$ is assigned to junction site $j \in J$ without an ES system, zero otherwise.
Z_j	Vector $(Z_{1j}, Z_{2j}, \ldots, Z_{\lvert I \rvert j})^T$.
E_j, $\hat{E}_j$	Interval length of the approximated uniform distribution of the pooled wind output faced by junction site j with and without an ES system, respectively.

outputs from i and k at site j. Matching the first and the second moments of $\mathbf{w}_{t,j}$, we obtain

$$\mu_j = \sum_{i\in I} Y_{ij}\mu_{ij}, \tag{12a}$$

$$\epsilon_j^2 = \sum_{i\in I} \epsilon_{ij}^2 Y_{ij} + \sum_{i,k\in I,\, i<k} 2\rho_{ikj}\epsilon_{ij}\epsilon_{kj} Y_{ij} Y_{kj} = Y_j^T \Sigma_j Y_j. \tag{12b}$$

Similarly, ϵ_j^2 and μ_j for site j having no ES system can be expressed by (12) with Y being replaced with Z. Note that it suffices to only know historical statistics of wind at individual farms as well as transmission per unit cost in order to compute the values of $\{\eta_{ij}, \mu_{ij}, \epsilon_{ij}, \rho_{ikj}\}$, $\forall\, i, k \in I,\ j \in J$.

For the lines between ES-equipped junction sites and the load center, the economic transmission capacity is given by (4). As discussed previously, by computing the dimensionless number θ_j we can predetermine which segment of the nonsmooth cost in (5) to be incorporated into our planning model before knowing the assignment of the wind farms. The candidate junction sites are thus categorized into the following two subsets based on θ_j:

$$\begin{cases} J_1 = \{ j \in J \mid \theta_j < \frac{1}{2} \}; \\ J_2 = J \backslash J_1. \end{cases}$$

With the above curtailment and pooling considerations as well as transmission capacity choices, the ES-transmission planning model is formulated as follows:

$$\begin{aligned}
\text{minimize} \quad & v_3(X, V, Y, Z, E, \hat{E}) \\
&= \sum_{j\in J}\Big[h_j X_j + \sum_{i\in I} g_{ij}(Y_{ij}+Z_{ij}) + g_j(X_j+V_j) \Big] \\
&\quad + \sum_{j\in J}\sum_{i\in I} \big[q_{ij}\mu_i + q_{ij}\big(\tfrac{1}{2}-\tfrac{1}{2}\eta_{ij}\big)\epsilon_i \big](Y_{ij}+Z_{ij}) \\
&\quad + \sum_{j\in J}\Big[q_j \Big(\sum_{i\in I} \mu_{ij} Z_{ij} + q_j\big(\tfrac{1}{2}-\tfrac{1}{2}\eta_j\big)\hat{E}_j \Big) \Big] \\
&\quad + \sum_{j\in J_1}\Big[q_j \sum_{i\in I} \mu_i Y_{ij} + q_j\big(\tfrac{1}{2}-\tfrac{1}{2}\theta_j\big)E_j \Big] \\
&\quad + \sum_{j\in J_2}\Big[q_j \sum_{i\in I} \mu_i Y_{ij} + \frac{p(1-\alpha\beta)}{8\delta} E_j \Big]
\end{aligned} \tag{13a}$$

$$\text{subject to} \quad \sqrt{Y_j^T \Sigma_j Y_j} \leqslant E_j, \quad \forall\, j \in J; \tag{13b}$$

$$\sqrt{Z_j^T \Sigma_j Z_j} \leqslant \hat{E}_j, \quad \forall\, j \in J; \tag{13c}$$

$$\sum_{j\in J}(Y_{ij}+Z_{ij}) = 1, \quad \forall\, i \in I; \tag{13d}$$

$$X_j + V_j \leqslant 1, \quad \forall\, j \in J; \tag{13e}$$

$$Y_{ij} \leqslant X_j, \quad \forall\, i \in I, \forall\, j \in J; \tag{13f}$$

$$Z_{ij} \leqslant V_j, \quad \forall\, i \in I, \forall\, j \in J; \tag{13g}$$

$$Y_{ij} = 0, \quad Z_{ij} = 0, \quad \forall\, (i,j) \in \{(i,j) \mid \eta_{ij} \geqslant 1\}; \tag{13h}$$

$$X_j = 0, \quad V_j = 0, \quad \forall\, j \in \{ j \mid \eta_j \geqslant 1\}; \tag{13i}$$

$$E_j, \hat{E}_j \geqslant 0, \quad \forall\, j \in J; \tag{13j}$$

$$X_j, V_j, Y_{ij}, Z_{ij} \in \{0,1\}, \quad \forall\, i \in I, \forall\, j \in J. \tag{13k}$$

In the above formulation, the objective (13a) is to minimize the total expected annual building and operating cost of the ES-transmission network for the given set of wind farms. The three terms in the first bracket are the fixed construction cost of ES systems, the fixed cost of transmission lines from the wind farms to the junction sites, and the fixed cost of transmission lines from the junction sites to the load center, respectively. The terms in the second and the third brackets are the variable costs of the ES-free transmission lines from the wind farms to the junction sites and from ES-free junction sites to the load center, respectively, according to (10). The uniform distribution parameter ϵ_j for junction site j with (without) an ES system is denoted by variable E_j ($\hat{E}_j$). The terms in the fourth and the fifth brackets are the variable costs of the economic lines from the two θ-categorized subsets of ES-equipped junction sites, J_1 and J_2, to the load center, with different cost expressions given by (5).

Constraints (13b) and (13c) define E_j and $\hat{E}_j$, respectively, based on Equation (12b). Constraints (13d) ensure that each wind farm in set I is assigned to one and only one junction site in set J. Constraints (13e) suggest that each candidate junction site can only get one of the three outcomes: to be selected and equipped with an ES system, to be selected without ES, or not to be selected. Constraints (13f) and (13g) require that a wind farm can only be assigned to a junction site that is selected. Constraints (13h) and (13i) exclude the potential assignment and junction site selection that would result in unprofitable transmission investment, as discussed in §4.2.

The above planning model is formulated as an MISOCP. The right-hand side of constraints (13b) and (13c) can be converted to the standard 2-norm form as $\|\tilde{\Sigma}_j Y_j\|_2 \leqslant E_j$ and $\|\tilde{\Sigma}_j Z_j\|_2 \leqslant \hat{E}_j$, respectively, $\forall\, j \in J$, where $\tilde{\Sigma}_j^T \tilde{\Sigma}_j = \Sigma_j$ since Σ_j is positive definite. Meanwhile, the objective function and all the other constraints are linear in the decision variables. In recent years, commercial software such as CPLEX have launched standard solvers for MISOCP. We will show that this planning model can be efficiently solved in a case of practical scale in §6. Please refer to Boyd and Vandenberghe (2004) and Alizadeh and Goldfarb (2001) for a comprehensive review of convex conic programs and MISOCP.

The above modeling process introduces two sources of inaccuracy. First, we apply uniform distribution approximations to wind outputs. As discussed in the numerical example in §4.3, this approximation at wind farms affects the sizing of the transmission lines that are upstream of the junction sites. Then we apply similar approximations to the wind outputs that are curtailed by the upstream ES-free lines and that are aggregated at the junction sites. These approximations affect the sizing of the transmission lines that are downstream of the junction sites. Second, planning model (13) tends to underestimate the transmission cost and thus oversize the lines. This is because the capacities of transmission lines upstream of the junction sites are evaluated based on (9), which overlooks transmission lines downstream of the junction sites.

However, we find that the model inaccuracy is well contained in practical settings, as briefly summarized in Table 3 for the cases with and without ES being colocated. In §EC.1 in the e-companion, each of those components of inaccuracy as well as the overall inaccuracy is quantified numerically and/or theoretically with detailed discussion. The theoretical studies also identify two more structural properties, namely, the curtailment independence of downstream transmission capacity and the decomposition of the joint optimization of upstream and downstream transmission capacities.

Planning model (13) can be extended in several ways to account for different investment considerations, such as maximum covering of wind farms. We omit such discussion for brevity, since those extensions are structurally similar.

Table 3. Overall model inaccuracy.

	ES-free (%)	ES-coupled (%)
Mean	8.25	6.42
Maximum	17.60	14.80

5. Capacitated Storage

In this section, we relax the assumption of uncapacitated ES and explicitly characterize the dependence of the energy overflow cost on ES and transmission capacities. Since the fluctuation of storage level is complicated and cannot be quantified in closed form, we instead derive a conservative estimate of the storage level distribution and then an upper bound for the optimal ES capacity. Moreover, we show in §§6 and EC.3 in the e-companion that this upper bound results in near-optimal expected total cost. The additional notation is summarized in Table 4.

5.1. Upper Bound of Energy Overflow

We choose relatively long periods (e.g., 1 day) with indices denoted as $\tau = 1, 2, \ldots$ and interval length as δ_b. In this way, the bulk wind output process, $\{\mathbf{w}_{b,\tau}\}$, is much less autocorrelated than the hourly process, due to diurnal cycles of wind speed (Thomann and Barfield 1988). We further assume that $\{\mathbf{w}_{b,\tau}\}$ is an independent and identical process. Such simplification causes underestimation of interperiod energy overflow loss when the storage level is nearly full. On the other hand, we implicitly assume that the wind output power within each interval τ is of constant value $\mathbf{w}_{b,\tau}/\delta_b$. Consequently, intraperiod energy overflow is overestimated, in that, in the long run, those real sample paths of wind output that are not constant but amount to the same bulk energy $\mathbf{w}_{b,\tau}$ within an interval result in more friction loss and thus less occurrences of $\mathbf{s}_\tau$ hitting S. These counteracting inaccuracies are further discussed in §§5.2 and EC.3 in the e-companion.

We use $\mathbf{s}_\tau$ to denote the storage level at the end of each interval τ, while assuming that the base storage capacity that accommodates periodic variation of storage level within each interval has already been captured by the fixed cost h. Our goal is to derive an economic storage capacity S in addition to the base capacity by obtaining the expected energy overflow as a function of S. As the first step, the

Table 4. Summary of additional notation.

	Systems parameters
δ_b	Length (in hours) of each time interval of independent bulk wind energy process.
$\mathbf{w}_{b,\tau}$	Random variable of the bulk energy a wind farm captures during interval $[\tau-1, \tau)$, where $\tau \in \mathbb{N}$ is the index of the intervals of length δ_b.
μ_b, ϵ_b	Mean and interval length of the approximated uniform distribution of $\mathbf{w}_{b,\tau}$, respectively.
$\mathbf{s}_\tau$	Storage level of an ES system at the end of interval τ.
$f_s(\cdot)$	Probability density function of $\mathbf{s}_\tau$.
$\mathbf{o}_\tau$	Energy overflow loss during interval τ.
	Decision variable
S	Maximum amount of potential energy that can be stored in an ES system.

transition of $\mathbf{s}_\tau$ is modeled as follows (Kim and Powell 2011):

$$\mathbf{s}_{\tau+1} = \begin{cases} S, \quad \text{if } \mathbf{s}_\tau + \alpha(\mathbf{w}_{b,\tau+1} - C\delta_b) \geqslant S; \\ \mathbf{s}_\tau + \alpha(\mathbf{w}_{b,\tau+1} - C\delta_b), \\ \quad \text{if } \mathbf{w}_{b,\tau+1} - C\delta_b > 0, \mathbf{s}_\tau \\ \qquad + \alpha(\mathbf{w}_{b,\tau+1} - C\delta_b) < S; \\ \mathbf{s}_\tau - (1/\beta)(C\delta_b - \mathbf{w}_{b,\tau+1}), \\ \quad \text{if } \mathbf{w}_{b,\tau+1} \leqslant C\delta_b < \beta \mathbf{s}_\tau + \mathbf{w}_{b,\tau+1}; \\ 0, \quad \text{if } C\delta_b \geqslant \beta \mathbf{s}_\tau + \mathbf{w}_{b,\tau+1}. \end{cases} \tag{14}$$

The four segments of the above piecewise linear function represent the four states of the storage level, respectively: fully charged, being charged, being discharged, and out of charge.

Deriving a closed-form expression of the probability density function $f_s(\cdot)$ for $\mathbf{s}_\tau$ is mathematically challenging, and hence we construct an approximation of $f_s(\cdot)$. We first make two assumptions: (1) The storage capacity S is large enough such that the probability of storage level switching from zero state to full state (or the other way around) within one interval is negligible. In fact, suppose that the ES capacity is small such that the above assumption is violated. Then the additional cost due to ES capacity and energy overflow becomes very small and dominated by the cost of transmission capacity. We will further examine this assumption in the numerical studies in §EC.3 in the e-companion. (2) We also assume that $f_s(s)$ is decreasing in s in the open interval $(0, S)$ when $C\delta_b$ is greater than or equal to the mean of $\mathbf{w}_{b,\tau}$. This assumption is realistic, since at certain storage levels s_τ, the probability of charge is greater than or equal to the probability of discharge. Also notice that, because of friction loss, any difference $\Delta S = |C\delta_b - \mathbf{w}_{b,\tau}|$ results in smaller magnitude of increase in s_τ when $C\delta_b > \mathbf{w}_{b,\tau}$ than the magnitude of decrease in s_τ when $C\delta_b < \mathbf{w}_{b,\tau}$ (see more detailed justification of this assumption in the proof of Proposition 1 in §EC.2.2 of the e-companion). Then we obtain the following proposition:

Proposition 1. *Assume* $C\delta_b \geqslant \mathbb{E}[\mathbf{w}_{b,\tau}]$, *and suppose* $\tilde{f}_s(s)$ *is an approximation of* $f_s(s)$ *such that* $\tilde{f}_s(s)$ *is constant in the open interval* $(0, S)$. *Then*
(i) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) \geqslant \mathbb{P}(\mathbf{s}_\tau = S)$;
(ii) $\tilde{\mathbb{E}}[\mathbf{o}_\tau] \geqslant \mathbb{E}[\mathbf{o}_\tau]$,
where $\tilde{\mathbb{P}}(\cdot)$ *and* $\tilde{\mathbb{E}}[\cdot]$ *denote probability and expectation with respect to* $\tilde{f}_s$, *respectively.*

Proposition 1(ii) establishes a sufficient condition to obtain an upper bound for the expected energy overflow $\mathbb{E}[\mathbf{o}_\tau]$. All proofs are given in §EC.2 in the e-companion. Intuitively, when $C\delta_b \geqslant \mathbb{E}[\mathbf{w}_{b,\tau}]$, the probability of an ES system being discharged is greater than or equal to its probability of being charged. As a result, $f_s(s)$ is decreasing in the open interval $(0, S)$ and it can be shown that $\tilde{f}_s(s)$ is greater than or equal to $f_s(s)$ when s is close to S. We also know that the expected overflow is nondecreasing in storage level once it is given. It thus can be shown that the approximated unconditioned overflow $\tilde{\mathbb{E}}[\mathbf{o}_\tau]$ is an upper bound for the real value.

To further obtain a closed-form upper bound for $\mathbb{E}[\mathbf{o}_\tau]$, we again approximate $\mathbf{w}_{b,\tau}$ by uniform distribution with mean μ_b and interval length ϵ_b, matching the first and the second moments of the real distribution of $\mathbf{w}_{b,\tau}$. The closed-form expression of $\tilde{f}_s(s)$ is given by the following proposition.

Proposition 2. *Assume* $\mathbf{w}_{b,\tau} \sim \text{Unif}(\mu_b - \epsilon_b/2, \mu_b + \epsilon_b/2)$. *Let* $A = (C\delta_b - \mu_b + \epsilon_b/2)^2/(2\beta(\mu_b + \epsilon_b/2 - C\delta_b))$ *and* $B = (\alpha(\mu_b + \epsilon_b/2 - C\delta_b)^2)/(2(C\delta_b - \mu_b + \epsilon_b/2))$. *Then*
(i) *for* $s \in (0, S)$, $\tilde{f}_s(s) \equiv f_s^c = 1/(A + S + B)$;
(ii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = 0) = A/(A + S + B)$;
(iii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) = B/(A + S + B)$.

Proposition 2 provides an analytical probability model for the distribution of storage level $\mathbf{s}_\tau$. Though it tends to overestimate the probability that $\mathbf{s}_\tau$ is close or equal to S, the model does capture the dependence of the storage level distribution on system parameters such as conversion efficiency and capacities of transmission and ES. In particular, Proposition 2 results in a simple analytical upper bound for the expected energy overflow, as stated in the following proposition:

Proposition 3. *Assume* $\mathbf{w}_{b,\tau} \sim \text{Unif}(\mu_b - \epsilon_b/2, \mu_b + \epsilon_b/2)$. *Then the expected energy overflow of each interval of length* δ_b *is bounded from above by* $(5\alpha/(24S))(\mu_b + \epsilon_b/2 - C\delta_b)^2$. *The derivatives of this upper bound are* $-(5\alpha/(24S^2))(\mu_b + \epsilon_b/2 - C\delta_b)^2$ *with respect to* S *and* $-(5\alpha\delta_b/(12S))(\mu_b + \epsilon_b/2 - C\delta_b)^2$ *with respect to* C.

An insight from Proposition 3 is that the upper bound of the expected overflow is more sensitive to C than to S. The marginal benefit in terms of overflow prevention shrinks quadratically in S. Therefore, transmission lines should be the dominant means over ES to reduce overflow. With the aid of Proposition 3, we next estimate ES capacity for a single and for multiple wind farms.

5.2. A Single Wind Farm

Our goal in this subsection is to find the economic trade-offs between transmission and ES capacities to minimize the total variable cost for a single wind farm. In particular, the variable ES cost consists of energy overflow loss and capacity cost. Following Proposition 3, an estimate of the annual variable ES cost, denoted by v_4, is as follows:

$$v_4(C, S) = \frac{5p\alpha}{24\delta\delta_b S}\left(\mu_b + \frac{\epsilon_b}{2} - C\delta_b\right)^2 + rS. \tag{15}$$

Minimizing the two terms on the right-hand side of (15) yields the upper bounds for economic storage capacity and the associated ES variable cost, both as functions of transmission capacity

$$S^*(C) = \sqrt{\frac{5p\alpha}{24r\delta\delta_b}}\left(\mu_b + \frac{\epsilon_b}{2} - C\delta_b\right), \tag{16}$$

$$v_4^*(C) = \sqrt{\frac{5pr\alpha}{6\delta\delta_b}}\left(\mu_b + \frac{\epsilon_b}{2} - C\delta_b\right). \tag{17}$$

Note that (16) and (17) are valid only if $C\delta_b \in (\mu_b, \mu_b + \epsilon_b/2)$. When $C\delta_b \geqslant \mu_b + \epsilon_b/2$, the transmission line is always able to deliver at least all the energy that is produced within the current interval by the end of the interval; as a result, no overflow occurs and $v_4^*(C) = 0$. Based on (3), (17), and the above discussion, the total variable cost is given by

$$v_5(C) = v_1(C) + v_4^*(C)$$
$$= \begin{cases} p\dfrac{1-\alpha\beta}{2\epsilon\delta}\left(\mu + \dfrac{\epsilon}{2} - C\right)^2 + qC \\ \quad + \sqrt{\dfrac{5pr\alpha}{6\delta\delta_b}}\left(\mu_b + \dfrac{\epsilon_b}{2} - C\delta_b\right), \\ \qquad \text{if } C \in \left[\dfrac{\mu_b}{\delta_b}, \dfrac{\mu_b + \epsilon_b/2}{\delta_b}\right); \\ p\dfrac{1-\alpha\beta}{2\epsilon\delta}\left(\mu + \dfrac{\epsilon}{2} - C\right)^2 + qC, \\ \qquad \text{if } C \in \left[\dfrac{\mu_b + \epsilon_b/2}{\delta_b}, \mu + \dfrac{\epsilon}{2}\right]. \end{cases} \tag{18}$$

In (18), $\mu_b/\delta_b = \mu$ and the economic value of C is within $[\mu, \mu + \epsilon/2]$. In addition, we assume $(\mu_b + \epsilon_b/2)/\delta_b < \mu + \epsilon/2$, since $\mathbf{w}_{b,\tau}$ aggregates hourly wind output $\mathbf{w}_t$ and is thus less variable. Function $v_5(C)$ is convex in C since it is the pointwise maximum of two convex quadratic functions of C. Therefore, the analytical expression of the economic transmission capacity C^* that minimizes (18), as well as the associated ES capacity S^* (from (16)) and the total variable cost v_5^*, can be obtained as summarized in Table 5 (In the table, recall $\theta = \delta q/p(1-\alpha\beta)$, and let $\hat{q} = q - \sqrt{5p\alpha r\delta_b/(6\delta)}$ and $\hat{\theta} = \delta\hat{q}/(p(1-\alpha\beta))$). These quantities can be viewed as generalization from (4) and (5), with ES capacity cost and overflow being considered.

In practice, the actual optimal cost may be greater than its theoretical upper bound v_5^* because of two sources of model inaccuracy—the uniform distribution approximation and the neglect of autocorrelation of the bulk wind output process. As for the latter, the resulting underestimation of interperiod overflow loss may overly compensate the overestimation of intraperiod overflow loss and ES capacity. Hence, v_5^* is closer to the actual optimal cost than theoretically expected. In the e-companion, §EC.3 presents detailed numerical studies on tightness and accuracy of the upper bound model. The studies verify that the cost gaps are bounded and v_5^* is not violated in most cases. Also, as an estimate to the actual optimal cost, v_5^* is more accurate than the cost lower bound developed in §4, as shown in Table 6. The table also shows that ES can potentially significantly bring down the cost of otherwise building an ES-free line.

5.3. Multiple Wind Farms

We next use the results in §5.2 to plan an ES-transmission network for multiple wind farms. However, these results can not be directly incorporated into the planning model. Unlike in the scenario of uncapacitated ES where we can predetermine which segment of the nonsmooth cost expression (5) to use in formulating planning model (13), the segmentation of v_5^* depends not only on wind-independent parameter θ, but also on wind characteristics $(\mu, \epsilon, \mu_d, \epsilon_d)$, which are not available for a candidate junction site before the assignment of the wind farms to the junction site is known. We hence resort to a heuristic outlined as follows.

1. For given sets of wind farms and candidate junction sites, solve planning model (13), which assumes uncapacitated ES. The model outputs an assignment of the wind farms

Table 5. Economic ES-transmission capacities and total variable costs.

Condition	C^*	S^*	v_5^*
$\dfrac{\mu_b+\epsilon_b/2}{\delta_b} < \mu + \epsilon\left(\frac{1}{2}-\theta\right)$	$\mu + \epsilon\left(\frac{1}{2}-\theta\right)$	0	$q\mu + \epsilon q\left(\frac{1}{2} - \frac{1}{2}\theta\right)$
$\dfrac{\mu_b+\epsilon_b/2}{\delta_b} \in \left[\mu + \epsilon\left(\frac{1}{2}-\theta\right), \mu + \epsilon\left(\frac{1}{2}-\hat{\theta}\right)\right)$	$\dfrac{\mu_b+\epsilon_b/2}{\delta_b}$	0	$q\left(\mu + \dfrac{\epsilon_b}{2\delta_b}\right) + p\dfrac{1-\alpha\beta}{8\epsilon\delta}\left(\epsilon - \dfrac{\epsilon_b}{\delta_b}\right)^2$
$\dfrac{\mu_b+\epsilon_b/2}{\delta_b} \geqslant \mu + \epsilon\left(\frac{1}{2}-\hat{\theta}\right) > \mu$	$\mu + \epsilon\left(\frac{1}{2}-\hat{\theta}\right)$	$\sqrt{\dfrac{5p\alpha}{24r\delta\delta_b}}\left(\dfrac{\epsilon_b}{2} - \epsilon\delta_b\left(\frac{1}{2}-\hat{\theta}\right)\right)$	$q\mu + \epsilon\hat{q}\left(\frac{1}{2} - \frac{1}{2}\hat{\theta}\right) + \sqrt{\dfrac{5pr\alpha}{6\delta\delta_b}}\dfrac{\epsilon_b}{2}$
$\mu + \epsilon\left(\frac{1}{2}-\hat{\theta}\right) \leqslant \mu$ (i.e., $\hat{\theta} \geqslant \frac{1}{2}$)	μ	$\sqrt{\dfrac{5p\alpha}{6r\delta\delta_b}}\left(\dfrac{\epsilon_b}{4}\right)$	$q\mu + \epsilon\dfrac{p(1-\alpha\beta)}{8\delta} + \sqrt{\dfrac{5pr\alpha}{6\delta\delta_b}}\dfrac{\epsilon_b}{2}$

to the junction sites as well as the associated transmission capacities.

2. For each junction site j that is selected to build ES on, compute the expected variable cost $v^*_{5,j}$ using Table 5. If $v^*_{5,j}+h_j<q_j\mu_j+q_j(\frac{1}{2}-\frac{1}{2}\eta_j)\epsilon_j$, then building ES on site j is still economical, and we choose capacities of ES and transmission line from site j to the load center according to Table 5. Otherwise, we opt for an ES-free line with capacity given by (9).

Then we summarize the models developed in this paper in the following proposition:

Proposition 4. *The optimal cost of the ES-transmission network is bounded from below by the optimal objective value of planning model* (13), *and bounded from above by the total cost given by the heuristic in* §5.3.

5.4. The Value of ES

All the above analysis suggests an important insight to the planners: *even small ES saves big, but the marginal value of ES diminishes fast*. Compared with the ES-free scenario, the combination of an economic line and an ES system has dual effects. The positive effect is that part of the investment in transmission capacity can be salvaged by the ES system, which accommodates short-period fluctuations of the wind output; the negative effect is the incurred energy loss because of friction and overflow. However, these dual effects respond to the size of the ES system differently. A relatively small capacity of ES should be adequate to achieve the positive effect. In contrast, additional ES capacity has much diminished value because it does not help reduce friction loss and is mainly used to hedge overflow. Moreover, Proposition 3 has suggested that the marginal benefit in terms of overflow prevention shrinks quadratically in S. As a result, building even more ES capacity would be less cost-efficient than increasing transmission capacity.

6. Computational Results and Insights

6.1. A Case Study

This section presents numerical studies to demonstrate the effectiveness of the proposed planning procedure and insights concerning technology impact and layout robustness.

Our first set of experiments generate ES-transmission network designs for potential wind farms near Billings, Montana. We solve planning model (13) to obtain network layouts as well as lower-bound total costs. Then the heuristic in §5.3 is applied to output upper-bound total costs as well as the associated transmission and ES capacities. The detailed settings of the experiments are described in §EC.4.1 in the e-companion. The experiments are repeated with different numbers of potential wind farms to be covered. Table 7 shows the costs and computational times of solving model (13). A network layout for the 24 wind farms is depicted in Figure 3(a). The correlation between wind outputs from those farms is shown in Figure 3(b).

Table 7 shows that the gap between the lower and the upper bounds of the total expected costs is smaller than that implied in Table 6. This is mainly because the additional fixed ES and transmission costs dilute the share of the cost that ES capacity and expected overflow account for. Also notice that, in most cases, the cost gap increases as we expand the network by adding distant wind farms. Again, this is mainly because the average transmission capacity decreases

Figure 3. (Color online) (a) Deployment of transmission lines and ES systems for 24 wind farms. (b) Correlation coefficients between wind outputs of the 24 wind farms.

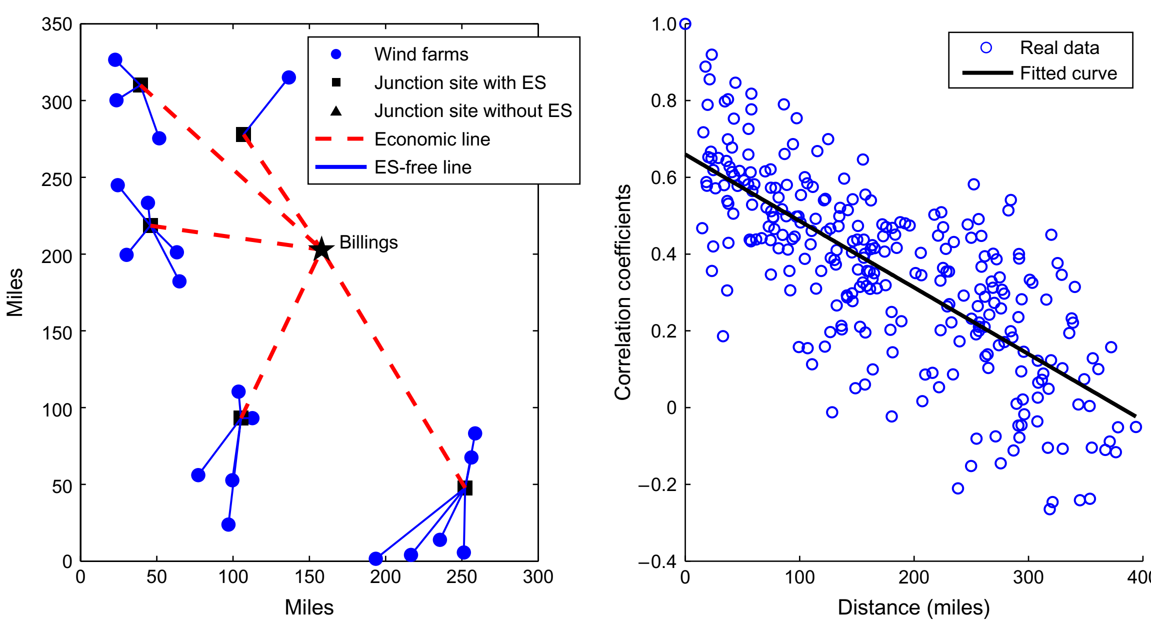


Table 6. Average cost gaps between the upper and the lower bounds, between the upper bound and the optimal value and between the upper bound and the cost in the ES-free scenario.

Distance (mile)	UB–LB (%)	|UB–Opt.| (%)	ES-Free–UB (%)
200	23.60	6.60	89.7
120	20.20	5.17	81.5
50	3.09	6.78	31.8

Table 7. Costs and computational times of the ES-transmission networks.

Wind farms	Candidate sites	LB cost ($)	UB cost ($)	Cost gap (%)	Time (sec)
6	12	1.9556×10^8	2.1136×10^8	8.08	0.6080
12	18	3.9703×10^8	4.2268×10^8	6.46	1.0264
18	28	6.2031×10^8	6.7153×10^8	8.26	2.0009
24	28	8.6438×10^9	9.4643×10^9	9.49	2.5876

in transmission distance; as a result, more ES capacity has to be added, which is not captured in the lower-bound cost. Finally, the computational times in all the instances are in magnitude of seconds. We thus conclude that, under the assumed settings of the ES-transmission network, our model is computationally efficient to solve and is able to output near-optimal ES-transmission deployment.

6.2. Technological Considerations

How do the forms of ES technologies and their advancements impact cost savings in power infrastructure planning? Various ES technologies will continue to be competing in the foreseeable future (Schoenung and Hassenzahl 2003). These technologies distinguish themselves from each other by storing potential energy in different ways with different cost-efficiency parameters. For example, pumped hydro storage systems have high upfront cost, low per-unit capital cost, and relatively high conversion efficiency (75%–78%), lead-acid batteries have high capital cost with high conversion efficiency, whereas compressed air energy storage systems feature very low capital cost with low conversion efficiency (Steward et al. 2009). As characterized in our models, these technological aspects impact the need for ES via different mechanisms. Furthermore, as these technologies advance and diffuse, the ES capacity cost and conversion efficiency remain uncertain to a large extent.

Our second set of experiments evaluate how the cost-savings respond to different values of these parameters. Our major finding is that the R&D priority should be given to addressing the bottleneck cost factor in order to maximize cost savings. For example, more investment is desirable in improving conversion efficiency in the case of compressed air energy storage systems, and in reducing the per-unit ES capacity cost for lead-acid batteries. More detailed analysis is omitted here for brevity and is available in §EC.4.2 in the e-companion.

6.3. Layout Robustness

In this subsection, we address the last issue with our model. Namely, is our planning model robust in the cost-minimizing sense? An ES-transmission network has to be planned in the presence of uncertainties, such as FIT rates adjustments, technology advancements, building material cost fluctuations, climate changes, etc. After infrastructure is built, it is still possible to locally adjust ES and transmission capacities. However, it will be too cost prohibitive to change the network layout as those uncertainties evolve over time. Therefore, a layout that is cost-efficient for a wide range of system parameter values is desirable.

The optimal layout may change at two levels. The topology-level reconfiguration affects the assignment of wind farms to junction sites. The switch over between installing and not installing an ES system at a junction site, and subsequently between building an economic and an ES-free transmission line downstream of it, is at the junction level.

We solve multiple instances of planning model (13) for cost-efficient layouts based on the settings of the case study in §6.1, perturbing parameters one at a time. Table 8 summarizes how the number of junction sites and the number of deployed ES systems change accordingly. These two numbers represent the layout changes at the topology level and junction level, respectively. The baseline layout has five ES-coupled junction sites, as depicted in Figure 3(a).

Table 8 shows that the baseline layout is remarkably robust. Set from 50% to 200% of the baseline value, almost no single parameter alone can affect the layout, except that one ES system is removed when either the variable transmission capital cost or the mean distance is halved. Further parameter deviations result in more significant layout changes, but are very unlikely to occur.

Table 8. The number of junction sites and the number of ES systems (in parentheses), as a parameter varies from 10% to 500% of its baseline value or the conversion efficiency varies from 0.40 to 0.99 (the N/A entries correspond to the cases where the infrastructure is not economically feasible).

Perturbed parameters	10%	50%	200%	500%
FIT rate (p)	N/A	5(5)	5(5)	5(4)
Fixed ES capacity cost (h)	5(5)	5(5)	5(5)	5(4)
Variable transmission capacity cost (a)	5(0)	5(4)	5(5)	N/A
Fixed transmission capacity cost (g)	8(8)	5(5)	5(5)	5(5)
Mean distance to the load center	5(0)	5(4)	5(5)	N/A
Perturbed parameter	0.40	0.60	0.80	0.99
ES round-trip conversion efficiency ($\alpha\beta$)	5(4)	5(5)	5(5)	5(5)

The above results suggest that the cost-efficient layout of the ES-transmission network is robust against gross misestimation of model parameters. Two reasons may explain this robustness. First, the whole system intricately depends on multiple parameters. Impact of misestimation of any one of them is substantially mitigated by the others. Even if more than one cost factor deviates, their influences upon layout is more likely to counteract than reinforce each other, with the fixed and variable transmission capacity costs being an example. Second, most of the two-echelon facility location models in the literature (for example, Shen et al. 2003, Cachon 2014) assume that the upstream facilities to be deployed are disjoint. In contrast, in our problem, all the selected junction sites are connected to a given load center. As a result, the favored junction sites not only cluster nearby wind farms, but also tend to be close to the given load center. This centrality of the junction sites substantially enhances the layout robustness.

7. Conclusion

In this paper, we study the problem of planning economic energy storage systems and transmission lines for distributed wind resources. Under the FIT policy instrument, the ES operating policy is to store surplus energy that exceeds the rated transmission capacity and release it later when the transmission line becomes available for additional loads. Although saving transmission capacity cost, operating ES systems incurs both energy friction loss and overflow loss over time, in addition to its building cost. We develop models to characterize these trade-offs and determine the sites and the sizes of ES systems as well as the associated topology and capacity of the transmission network. In our first model, under the assumption that ES systems have sufficient energy capacity, we derive optimal transmission line capacities for a single wind farm with and without ES being colocated, respectively. Then we incorporate these quantities to formulate a location model as an MISOCP to determine the ES-transmission network. Our second model addresses the ES sizing problem. We derive an upper bound of the expected energy overflow, from which an overestimated economic storage size as well as the associated transmission capacity is obtained. These two models provide the lower and the upper bounds of the total expected cost. The case study, as well as the analysis of model inaccuracy in the e-companion, demonstrates that our deployment scheme is near optimal.

Our model and analysis lead to several findings. (1) First, although utilization of ES systems saves a significant amount of transmission cost, the marginal value of adding ES capacity diminishes faster than the marginal value of adding transmission capacity. This is because a relatively small amount of ES capacity is sufficient for smoothing short-term fluctuations of wind outputs; but adding more ES capacity does not help reduce energy friction loss and is inefficient in reducing energy overflow loss compared with adding transmission capacity. Therefore, it is cost-efficient to install small-sized ES systems on electricity junction sites. (2) Advancements in ES technologies can further bring down the total cost of an ES-transmission network by improving round-trip energy conversion efficiency and reducing upfront cost and per-unit capacity cost of ES. Depending on the form of the ES technology, the priority of R&D resources should be given to addressing the bottleneck cost factor. For example, the bottleneck cost factor is the round-trip conversion efficiency for compressed air energy storage systems, and is the per-unit capacity cost for lead-acid batteries. (3) The layout of the ES-transmission network is robust against FIT rates adjustments, technology advancements, building material cost fluctuations, climate changes, etc. This is partly because the system relies on multiple parameters and partly because of the centrality of the favored junction sites toward the given load center in our problem setting.

This paper is the first attempt to provide tractable methods to determine both the layout and capacity of an ES-transmission network to tap distributed wind resources. The work adopts some core ideas from the field of supply chain design. We believe that more contribution to planning renewable energy systems can be made by the supply chain research community. Two possible extensions to this work can be beneficial. Firstly, despite the prevalence of the FIT policy instrument in the initial phase, wind energy is expected to enter the energy biding market when it accounts for a considerable share of the total energy production in the future. Under such a scenario, the process of wind energy prices are stochastic, and more ES capacity is desired to store energy and shift its delivery to high-price hours. The question then arises: how to make adaptive budgeting decisions on investment in transmission lines and ES systems given the evolving wind energy market (i.e., from FIT based to bid based)? Secondly, although the radial topology with a single load center is a realistic setting for the initial stage of wind energy development and captures most of the trade-offs in the scenario of covering multiple wind farms, it will be more interesting (and much more challenging) to develop models that address the complexity where the grid is meshed and multiple load centers are present.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2015.1444.

Acknowledgments

The authors thank Eddie Anderson (the area editor), the associate editor, and two anonymous referees for their valuable comments. The authors are also in great debt to Shmuel Oren of the Industrial Engineering and Operations Research Department, Duncan Callaway of the Energy and Resources Group, both at UC Berkeley, and Zechun Hu of the Electrical Engineering Department at Tsinghua University, for the helpful discussions on various aspects of this paper. The authors would like to thank National Renewable Energy Laboratory for providing data sets of wind power. This research was partially supported by the National Science Foundation [Grant CMMI 1265671] and the National Science Foundation of China [Grants 71210002 and 71332005].

References

Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS (1999) Wind speed and power forecasting based on spatial correlation models. *Energy Conversion, IEEE Trans.* 14(3):836–842.

Alizadeh F, Goldfarb D (2001) Second-order cone programming. *Math. Programming* 95:3–51.

Alizamir S, de Véricourt F, Sun P (2012) Efficient feed-in-tariff policies for renewable energy technologies. Working paper, Yale University, New Haven, CT.

Atamtürk A, Berenguer G, Shen ZJM (2012) A conic integer programming approach to stochastic joint location-inventory problems. *Oper. Res.* 60(2):366–381.

AWC (2014) Atlantic wind connection project. Accessed January 15, http://atlanticwindconnection.com/awc-projects/atlantic-wind-connection.

Baringo L, Conejo AJ (2012) Transmission and wind power investment. *Power Systems, IEEE Trans.* 27(2):885–893.

Boyd S, Vandenberghe L (2004) *Convex Optimization* (Cambridge University Press, New York).

Cachon GP (2014) Retail store density and the cost of greenhouse gas emissions. *Management Sci.* 60(8):1907–1925.

California ISO (2013) 2012–2013 ISO transmission plan. Accessed January 14, 2014, https://www.caiso.com/planning/Pages/TransmissionPlanning/2012-2013TransmissionPlanningProcess.aspx.

Denholm P, Sioshansi R (2009) The value of compressed air energy storage with wind in transmission-constrained electric power systems. *Energy Policy* 37(8):3149–3158.

Department of Energy (2008) 20% Wind energy by 2030—Increasing wind energy's contribution to U.S. electricity supply. Report, U.S. Department of Energy. Accessed June 5, 2013, http://www.nrel.gov/docs/fy08osti/41869.pdf.

Electricity Advisory Committee (2008) Bottling electricity: Storage as a strategic tool for managing variability and capacity concerns in the modern grid. Report, Electricity Advisory Committee. Accessed June 5, 2013, http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/final-energy-storage_12-16-08.pdf.

Energy Information Administration (2014) Annual energy outlook 2014. Report, U.S. Energy Information Administration. Accessed January 23, 2015, http://www.eia.gov/forecasts/aeo/pdf/0383%282014%29.pdf.

Elliott D, Schwartz M, Scott G, Haymes S, Heimiller D, George R (2001) Wind energy resource atlas of mongolia. Technical report, National Renewable Energy Lab., Golden, CO.

Elliott DL, Wendell LL, Gower GL (1991) An assessment of the available windy land area and wind energy potential in the contiguous United States. Technical report, Pacific Northwest Lab., Richland, WA.

EPRI-DOE (2003) EPRI-DOE handbook of energy storage for transmission and distribution applications. Electric Power Research Institute and U.S. Department of Energy.

NREL (2008) Western wind power resources database. Report, National Renewable Energy Laboratory.

Greenblatt JB, Succar S, Denkenberger DC, Williams RH, Socolow RH (2007) Baseload wind energy: Modeling the competition between gas turbines and compressed air energy storage for supplemental generation. *Energy Policy* 35(3):1474–1492.

Greenpeace (2014) Clicking clean: How companies are creating the green internet. Report, Greenpeace Inc, Washington, DC.

Hemmati R, Hooshmand RA, Khodabakhshian A (2013) State-of-the-art of transmission expansion planning: Comprehensive review. *Renewable Sustainable Energy Rev.* 23:312–319.

Kaltenbach JC, Peschon J, Gehrig EH (1970) A mathematical optimization technique for the expansion of electric power transmission systems. *Power Apparatus and Systems, IEEE Trans.* 89(1):113–119.

Kim JH, Powell WB (2011) Optimal energy commitments with storage and intermittent supply. *Oper. Res.* 59(6):1347–1360.

Kim KJ, Park YM, Lee KY (1988) Optimal long term transmission expansion planning based on maximum principle. *Power Systems, IEEE Trans.* 3(4):1494–1501.

Latorre G, Cruz RD, Areiza JM, Villegas A (2003) Classification of publications and models on transmission expansion planning. *Power Systems, IEEE Trans.* 18(2):938–946.

Mak HY, Rong Y, Shen ZJM (2013) Infrastructure planning for electric vehicles with battery swapping. *Management Sci.* 59(7):1557–1575.

Mason T, Curry T, Wilson D (2012a) Capital costs for transmission and substations: Recommendations for WECC transmission expansion planning. Report, Black and Veatch, . Accessed May 6, 2015, https://www.wecc.biz/Reliability/1210_BV_WECC_TransCostReport_Final.pdf.

Mason T, Curry T, Wilson D (2012b) Transmission capital cost for WECC-TEPPC. Accessed April 4, 2015.

Mendonça M, Jacobs D, Sovacool BK (2009) *Powering the Green Economy: The Feed-in Tariff Handbook* (Earthscan, London).

Moeini-Aghtaie M, Abbaspour A, Fotuhi-Firuzabad M (2012) Incorporating large-scale distant wind farms in probabilistic transmission expansion planning—part i: Theory and algorithm. *Power Systems, IEEE Trans.* 27(3):1585–1593.

Natarajan K, Sim M, Uichanco J (2010) Tractable robust expected utility and risk models for portfolio optimization. *Math. Finance* 20(4):695–731.

Steward D, Saur G, Penev, M, Ramsden T (2009) Lifecycle cost analysis of hydrogen versus other technologies for electrical energy storage. Technical Report, National Renewable Energy Laboratory. Accessed February 2, 2014, http://www.nrel.gov/docs/fy10osti/46719.pdf.

Oh H (2011) Optimal planning to include storage devices in power systems. *Power Systems, IEEE Trans.* 26(3):1118–1128.

REN21 (2011) Renewable energy policy network for the 21st century. Renewables 2011 global status report. Accessed May 25, 2013, http://germanwatch.org/klima/gsr2011.pdf.

Rogers J, Fink S, Porter K (2010) Examples of wind energy curtailment practices. Subcontract report, National Renewable Energy Laboratory. Accessed June 5, 2013, http://www.nrel.gov/docs/fy10osti/48737.pdf.

Schoenung SM, Hassenzahl WV (2003) Long- vs. short-term energy storage technologies analysis: A life-cycle cost study. A study for the DOE energy storage systems program. Report, Sandia National Laboratories. Accessed May 27, 2013, http://prod.sandia.gov/techlib/access-control.cgi/2003/032783.pdf.

Shen ZJM, Coullard C, Daskin MS (2003) A joint location-inventory model. *Transportation Sci.* 37(1):40–55.

Talinli I, Topuz E, Aydin E, Kabakci S (2011) A holistic approach for wind farm site selection by using FAHP. Suvire GO, ed. *Wind Farm—Technical Regulations, Potential Estimation and Siting Assessment, InTech*. Accessed September 17, 2013, http://www.intechopen.com/books/wind-farm-technical-regulations-potential-estimation-and-siting-assessment.

Taylor JA, Hover FS (2013) Conic AC transmission system planning. *Power Systems, IEEE Trans.* 28(2):952–959.

Thomann GC, Barfield MJ (1988) The time variation of wind speeds and wind farm power output in Kansas. *Energy Conversion, IEEE Trans.* 3(1):44–49.

Zhang F, Hu Z, Song Y (2013) Mixed-integer linear model for transmission expansion planning with line losses and energy storage systems. *IET Generation, Transmission Distribution* 7(8):919–928.

Wei Qi is an assistant professor in the Department of Industrial Engineering at Tsinghua University. His research interests include design and management of supply chains and energy systems, sustainable operations, and data analytics.

Yong Liang is an assistant professor in the Department of Management Science and Engineering of the School of Economics and Management at Tsinghua University. His research interests span the areas of supply chain design, energy economics and operations, data center infrastructure provisioning optimization, and healthcare.

Zuo-Jun (Max) Shen is the Chancellor's Professor in the Department of Industrial Engineering and Operations Research and the Department of Civil and Environmental Engineering at UC Berkeley. He is also an honorary professor at Tsinghua University. He has been active in the following research areas: integrated supply chain design and management, market mechanism design, applied optimization, and decision making with limited information.