Submitted to *Operations Research*
manuscript OPRE-2014-07-437.R2

Joint Planning of Energy Storage and Transmission for Wind Energy Generation

Wei Qi
Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720

Yong Liang
School of Economics and Management, Tsinghua University, Beijing 100084, China

Zuo-Jun Max Shen
Department of Industrial Engineering and Operations Research and Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720

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 grid, and mitigate the intermittency of wind outputs. In this paper, we propose models of transmission network planning with co-location 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 (FIT) policy instrument. We first formulate a location model as a mixed integer second-order-conic program (MISOCP) 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.

Key words: Facility Location; Wind Energy; Energy Storage; Infrastructure Planning

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 renewable energy to account for 15% of its total consumption by 2020 (REN21 (2011)). In the mean time, leading companies in the 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); Elliott et al. (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 15MW for 6-8% of the time as of 2010 due to transmission constraints (Rogers et al. (2010)).

We try to address these challenges by proposing models of transmission network planning with co-location 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 co-locating 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 120MW data center with 100 percent 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 (Mendonca 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 notation and 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, section 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 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) and 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) and 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 due to charge/discharge friction loss and overflow loss of energy. Meanwhile, while 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. As 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 **??**. 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.

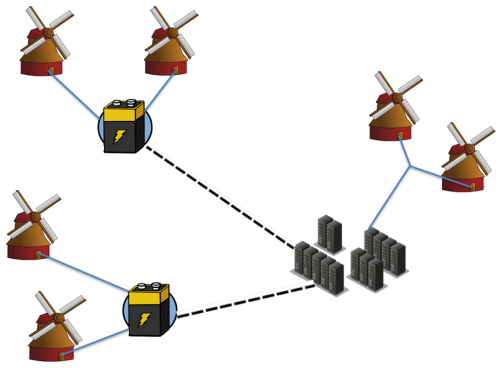

Figure 1 A radial ES-transmission network with economic (dashed) and ES-free (solid) lines.

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 non-perfect roundtrip conversion efficiency 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.

As 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 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 co-located 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 Section 5 we will show that this uncapacitated case provides a reasonable approximation to the case with capacity limts. Table 1 summarizes the notation for the following three subsections. 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 $\delta = 1\text{hr} = \frac{1}{24\times 365}\text{yr}$ 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 $\delta = 1\text{hr} = \frac{1}{24\times 365}\text{yr}$.

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 co-locate the ES system with the wind farm to avoid the cost of building 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 tradeoff, 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: $\mathbf{w}_t \sim \text{uniform}(\mu - \frac{\epsilon}{2}, \mu + \frac{\epsilon}{2})$. 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 Subsection 4.3 and in Subsection EC.1.1 in the e-companion. Note that the wind output process $\{\mathbf{w}_t\}$ can be auto-correlated and non-stationary. When $C \geq \mu - \frac{\epsilon}{2}$, the expected energy loss in $[t-1,t)$ is given by:

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

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) and Kim et al. (1988)). Recent empirical evaluation (Mason et al. (2012a) and Mason et al. (2012b)) 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:

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

It can be verified that $C = (\mu + \frac{\epsilon}{2}) - \frac{\epsilon\delta q}{p(1-\alpha\beta)}$ minimizes (3). We make an additional assumption that transmission line capacity should be greater than or equal to the average wind output power; otherwise the expected energy storage level could be infinite. Hence, the economic transmission capacity which minimizes the expected annual variable cost v_1 is given by:

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

where $\theta = \frac{\delta q}{p(1-\alpha\beta)}$.

The dimensionless number θ captures the cost associated with building transmission capacity. For example, if q is large due to 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 $\theta \geq \frac{1}{2}$, 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 pre-determine which segment of the following non-smooth 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:

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

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 co-located. In this case, the optimal transmission capacity can be directly expressed as the optimal quantile in the classic newsvendor model with stockout cost $(p - q\delta)$ and inventory holding cost $q\delta$. An explicit derivation resembles the steps in Subsection 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 $C \geq \mu - \frac{\epsilon}{2}$, the expected curtailment loss is given by:

$$\begin{aligned} \mathbb{E}[\mathbf{l}_t] &= \textstyle\int_C^{\mu+\epsilon/2} (w_t - C)\frac{1}{\epsilon} dw_t \\ &= \tfrac{1}{2\epsilon}(\mu + \tfrac{\epsilon}{2} - C)^2. \end{aligned} \tag{7}$$

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

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

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

$$
\begin{aligned}
C^* &= \begin{cases} \arg\min v_2(C), & \text{if } \eta < 1; \\ 0, & \text{otherwise;} \end{cases} \\
&= \begin{cases} \mu + (\frac{1}{2} - \eta)\epsilon, & \text{if } \eta < 1; \\ 0, & \text{otherwise,} \end{cases}
\end{aligned} \tag{9}
$$

where $\eta = \frac{\delta q}{p}$.

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 level of line capacity restriction. When $\eta \geq 1$, 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:

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

.

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 $p = \$0.08$/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 $a = \$1$/kW-mile.

Figure 2 (a) and (b) are the area plots of the total variable cost with and without ES being co-located, respectively. We use wind output data of a modeled site from EPRI-DOE (2008) (The site ID is 24648). 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 $l \geq 98$ miles, $C^* = \mu$, and hence the friction cost

reaches standstill while 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 $l \leq 50$ miles and 0.22% (1.09%) for $l > 50$ 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 the following subsection.

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 co-located.

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

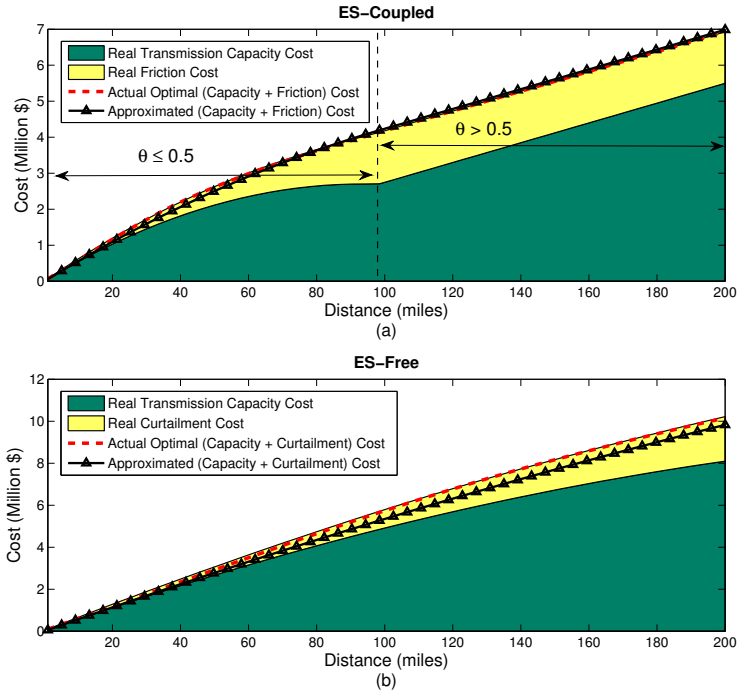

Figure 2 Variable costs of a single farm (a) with and (b) without ES.

in Subsections 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 Subsection EC.2.1 in the e-companion):

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

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

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., $\mathbf{w_{t,j}} \sim \text{Unif}(\mu_j - \frac{\epsilon_j}{2}, \mu_j + \frac{\epsilon_j}{2})$. Let Σ_j be the covariance matrix of wind outputs from the wind farms in I and aggregated at

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	1 if an ES system is built on site $j \in J$, 0 otherwise.
V_j	1 if site $j \in J$ is selected as a junction site with no ES system, 0 otherwise.
Y_{ij}	1 if wind farm $i \in I$ is assigned to junction site j with an ES system, 0 otherwise.
Y_j	vector $(Y_{1j}, Y_{2j}, ..., Y_{\|I\|j})^T$.
Z_{ij}	1 if wind farm $i \in I$ is assigned to junction site $j \in J$ without an ES system, 0 otherwise.
Z_j	vector $(Z_{1j}, Z_{2j}, ..., Z_{\|I\|j})^T$.
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.

junction site j. Each entry $\Sigma_{ikj} = \epsilon_{ij}\rho_{ikj}\epsilon_{kj}$, where ρ_{ikj} is the correlation coefficient of the curtailed wind 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 pre-determine which segment of the non-smooth 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:

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

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

$$\begin{aligned}
\text{minimize} \quad & v_3(X,V,Y,Z,E,\hat{E}) = \sum_{j\in J}[h_j X_j + \sum_{i\in I} g_{ij}(Y_{ij}+Z_{ij}) + g_j(X_j+V_j)] \\
& + \sum_{j\in J}\sum_{i\in I}[q_{ij}\mu_i + q_{ij}(\frac{1}{2}-\frac{1}{2}\eta_{ij})\epsilon_i](Y_{ij}+Z_{ij}) \\
& + \sum_{j\in J}[q_j(\sum_{i\in I}\mu_{ij}Z_{ij} + q_j(\frac{1}{2}-\frac{1}{2}\eta_j)\hat{E}_j)] \\
& + \sum_{j\in J_1}[q_j\sum_{i\in I}\mu_i Y_{ij} + q_j(\frac{1}{2}-\frac{1}{2}\theta_j)E_j] + \sum_{j\in J_2}[q_j\sum_{i\in I}\mu_i Y_{ij} + \frac{p(1-\alpha\beta)}{8\delta}E_j] && \text{(13a)} \\
\text{subject to} \quad & \sqrt{Y_j^T\Sigma_j Y_j} \le E_j \quad \forall j \in J && \text{(13b)} \\
& \sqrt{Z_j^T\Sigma_j Z_j} \le \hat{E}_j \quad \forall j \in J && \text{(13c)} \\
& \sum_{j\in J}(Y_{ij}+Z_{ij}) = 1 \quad \forall i \in I && \text{(13d)} \\
& X_j + V_j \le 1 \quad \forall j \in J && \text{(13e)} \\
& Y_{ij} \le X_j \quad \forall i \in I, \forall j \in J && \text{(13f)} \\
& Z_{ij} \le V_j \quad \forall i \in I, \forall j \in J && \text{(13g)}
\end{aligned}$$

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

$$X_j = 0, V_j = 0 \quad \forall j \in \{j | \eta_j \geq 1\} \tag{13i}$$

$$E_j, \hat{E}_j \geq 0 \qquad \forall j \in J \tag{13j}$$

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

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 Subsection 4.2.

The above planning model is formulated as a mixed-integer second order cone program (MISOCP). The right-hand side of constraints (13b) and (13c) can be converted to the standard 2-norm form as $\| \tilde{\Sigma}_j Y_j \|_2 \leq E_j$ and $\| \tilde{\Sigma}_j Z_j \|_2 \leq \hat{E}_j$, 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. Recent years commercial softwares 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 Section 6. Please refer to Boyd and Vandenberghe (2004) and Alizadeh and Goldfarb (2001) for 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 Subsection 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 co-located. In Section 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.

Table 3 Overall model inaccuracy.

	ES-Free	ES-Coupled
Mean	8.25%	6.42%
Maximum	17.6%	14.8%

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

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.

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 Sections 6 and EC.3 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 $\tau = 1, 2, ...$ and interval length as δ_b. In this way, the bulk wind output process, $\{\mathbf{w}_{b,\tau}\}$, is much less auto-correlated 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 inter-period 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 $\frac{\mathbf{w}_{b,\tau}}{\delta_b}$. Consequently, intra-period 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 the next subsection and Section 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 transition of $\mathbf{s}_\tau$ is modeled as follows (Kim and Powell (2011)):

$$\mathbf{s}_{\tau+\mathbf{1}} = \begin{cases} S, & \text{if } \mathbf{s}_\tau + \alpha(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} - C\delta_b) \geq S; \\ \mathbf{s}_\tau + \alpha(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} - C\delta_b), & \text{if } \mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} - C\delta_b > 0, \mathbf{s}_\tau + \alpha(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} - C\delta_b) < S; \\ \mathbf{s}_\tau - \frac{1}{\beta}(C\delta_b - \mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}), & \text{if } \mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} \leq C\delta_b < \beta\mathbf{s}_\tau + \mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}; \\ 0, & \text{if } C\delta_b \geq \beta\mathbf{s}_\tau + \mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}. \end{cases} \tag{14}$$

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 Section

EC.3. (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 level s_τ, the probability of charge is greater than or equal to the probability of discharge. Also notice that, due to 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 Subsection EC.2.2). Then we obtain the following proposition:

PROPOSITION 1. *Assume $C\delta_b \geq \mathbb{E}[\mathbf{w}_{b,\tau}]$, 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) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) \geq \mathbb{P}(\mathbf{s}_\tau = S)$;

(ii) $\tilde{\mathbb{E}}[\mathbf{o}_\tau] \geq \mathbb{E}[\mathbf{o}_\tau]$,

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 Section EC.2. Intuitively, when $C\delta_b \geq \mathbb{E}[\mathbf{w}_{b,\tau}]$, 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 non-decreasing 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 $\mathbf{w}_{b,\tau} \sim unif(\mu_b - \frac{\epsilon_b}{2}, \mu_b + \frac{\epsilon_b}{2})$. Let $A = \frac{(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)}$ and $B = \frac{\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2(C\delta_b - \mu_b + \frac{\epsilon_b}{2})}$. Then*

(i) For $s \in (0, S)$, $\tilde{f}_s(s) \equiv f_s^c = \frac{1}{A+S+B}$;

(ii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = 0) = \frac{A}{A+S+B}$;

(iii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) = \frac{B}{A+S+B}$.

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

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

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:

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

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

Note that (16) and (17) are valid only if $C\delta_b \in (\mu_b, \mu_b + \frac{\epsilon_b}{2})$. When $C\delta_b \geq \mu_b + \frac{\epsilon_b}{2}$, the transmission line is always able to deliver at least all the energy that are 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:

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

In (18), $\frac{\mu_b}{\delta_b} = \mu$ and the economic value of C is within $[\mu, \mu + \frac{\epsilon}{2}]$. In addition, we assume $\frac{\mu_b + \frac{\epsilon_b}{2}}{\delta_b} < \mu + \frac{\epsilon}{2}$, since $\mathbf{w}_{b,\tau}$ aggregates hourly wind output $\mathbf{w}_t$ and is thus less variable. $v_5(C)$ is convex in C since it is the point-wise 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 $\theta = \frac{\delta q}{p(1-\alpha\beta)}$, and let $\hat{q} = q - \sqrt{\frac{5p\alpha r\delta_b}{6\delta}}$ and $\hat{\theta} = \frac{\delta\hat{q}}{p(1-\alpha\beta)}$). These quantities can be viewed as generalization from (4) and (5), with ES capacity cost and overflow being considered.

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

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

In practice, the actual optimal cost may be greater than its theoretical upper bound v_5^* due to two sources of model inaccuracy - the uniform distribution approximation and the neglect of auto-correlation of the bulk wind output process. As for the latter, the resulting underestimation of inter-period overflow loss may overly compensate the overestimation of intra-period overflow loss

and ES capacity. Hence, v_5^* is closer to the actual optimal cost than theoretically expected. In the e-companion, Section 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 Section 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.

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	23.6%	6.6%	89.7%
120	20.2%	5.17%	81.5%
50	3.09%	6.78%	31.8%

5.3. Multiple Wind Farms

We next use the results in Subsection 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 pre-determine which segment of the non-smooth 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 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 Subsection 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 due to 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 subsection 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 Subsection 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 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

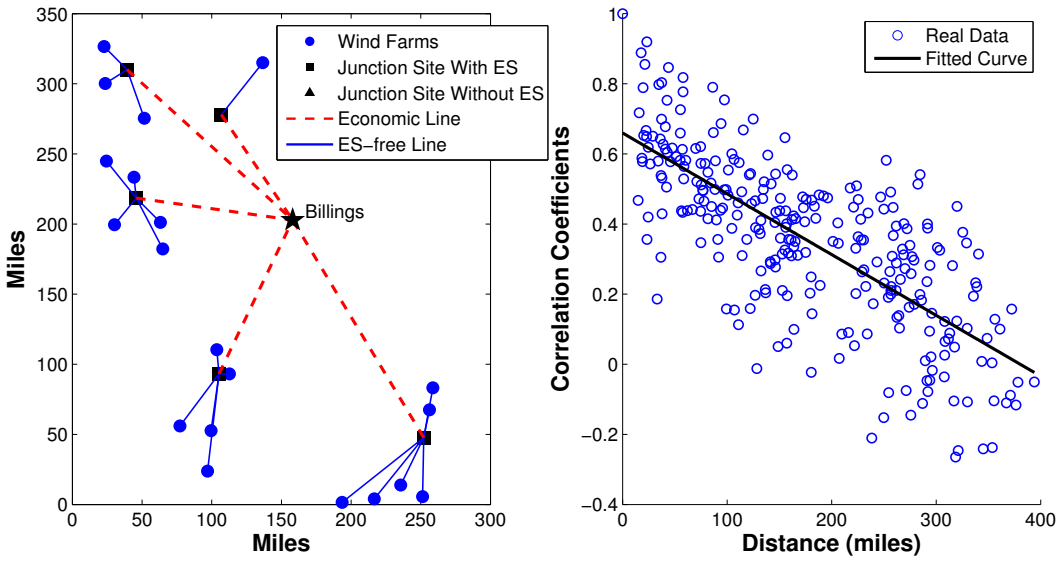

Figure 3 (a) Deployment of transmission lines and ES systems for 24 wind farms. (b) Correlation coefficients between wind outputs of the 24 wind farms.

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 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 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, while compressed air energy storage systems feature very low capital cost with low conversion efficiency (NREL (2012)). 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 Subsection 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. At the junction level is 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.

We solve multiple instances of planning model (13) for cost-efficient layouts based on the settings of the case study in Subsection 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- and junction- levels, 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.

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 factors deviate, their influences upon layout is more likely to counteract than reinforce each other, with the fixed and

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)

variable transmission capacity costs being an example. Second, most of the two-echelon facility location models in literature (for example, Shen et al. (2003) and 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. While 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 co-located, 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 provides 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, while utilization of ES systems saves 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 towards 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 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. In that case, AC power flows need to be carefully calculated.

Acknowledgments

The authors thank Eddie Anderson (the area editor), the associate editor, and two anonymous referees for their valuable comments; 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 helpful discussion on various aspects of this paper. The authors would like to thank NREL 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, 71332005.

References

Alexiadis, M.C., P.S. Dokopoulos, H.S. Sahsamanoglou. 1999. Wind speed and power forecasting based on spatial correlation models. *Energy Conversion, IEEE Transactions on* **14**(3) 836–842.

Alizadeh, F., D. Goldfarb. 2001. Second-order cone programming. *Mathematical Programming* **95** 3–51.

Alizamir, S., F. de Véricourt, P. Sun. 2012. Efficient feed-in-tariff policies for renewable energy technologies, *working paper* .

Atamtürk, A., G. Berenguer, Z.J.M. Shen. 2012. A conic integer programming approach to stochastic joint location-inventory problems. *Operations Research* **60**(2) 366–381.

AWC. 2014. Atlantic wind connection project. [Online; accessed on 15-January-2014].

Baringo, L., A.J. Conejo. 2012. Transmission and wind power investment. *Power Systems, IEEE Transactions on* **27**(2) 885–893.

Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, New York, NY, USA.

Cachon, G. P. 2014. Retail store density and the cost of greenhouse gas emissions. *Management Science* **60**(8) 1907–1925.

California ISO. 2013. 2012-2013 ISO transmission plan. [Revised draft; online; accessed on 14-Janurary-2014].

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

DOE. 2008. 20 percent wind energy by 2030 - increasing wind energy's contribution to U.S. electricity supply. US Department of Energy. [Online; accessed on 5-June-2013].

EAC. 2008. Bottling electricity: Storage as a strategic tool for managing variability and capacity concerns in the modern grid. Electricity Advisory Committee. [Online; accessed on 5-June-2013].

EIA. 2014. Annual energy outlook 2014. U.S. Energy Information Administration.

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

Elliott, D.L., L.L. Wendell, G.L. Gower. 1991. An assessment of the available windy land area and wind energy potential in the contiguous united states. Tech. rep., Pacific Northwest Lab., Richland, WA.

EPRI-DOE. 2003. Handbook of energy storage for transmission & distribution applications. US Department of Energy.

EPRI-DOE. 2008. Western wind power resources database. National Renewable Energy Laboratory.

Greenblatt, J.B., S. Succar, D.C. Denkenberger, R.H. Williams, R.H. Socolow. 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. Greenpeace Inc.

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

Kaltenbach, J.C., J. Peschon, E.H. Gehrig. 1970. A mathematical optimization technique for the expansion of electric power transmission systems. *Power Apparatus and Systems, IEEE Transactions on* **89**(1) 113–119.

Kim, J.H., W.B. Powell. 2011. Optimal energy commitments with storage and intermittent supply. *Operations Research* **59**(6) 1347–1360.

Kim, K.J., Y.M. Park, K.Y. Lee. 1988. Optimal long term transmission expansion planning based on maximum principle. *Power Systems, IEEE Transactions on* **3**(4) 1494–1501.

Latorre, G., R.D. Cruz, J.M. Areiza, A. Villegas. 2003. Classification of publications and models on transmission expansion planning. *Power Systems, IEEE Transactions on* **18**(2) 938–946.

Löfberg, J. 2004. Yalmip : a toolbox for modeling and optimization in MATLAB. *Proceedings of the CACSD Conference*. Taipei, Taiwan.

Mak, H.Y., Y. Rong, Z.J.M. Shen. 2013. Infrastructure planning for electric vehicles with battery swapping. *Management Science* **59**(7) 1557–1575.

Mason, T., T. Curry, D. Wilson. 2012a. Capital costs for transmission and substations: recommendations for WECC transmission expansion planning. *Black & Veatch* .

Mason, T., T. Curry, D. Wilson. 2012b. Transmission capital cost for WECC-TEPPC. [Online; accessed on 4-April-2015].

Mendonca, M., D. Jacobs, B. K. Sovacool. 2009. *Powering the green economy: The feed-in tariff handbook*. Earthscan.

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

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

NREL. 2012. Lifecycle cost analysis of hydrogen versus other technologies for electrical energy storage. National Renewable Energy Laboratory.

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

REN21. 2011. Renewables 2011 global status report. Renewable Energy Policy Network for the 21st Century.

Rogers, J., S. Fink, K. Porter. 2010. Examples of wind energy curtailment practices. *Subcontract Report NREL/SR-550* **48737**.

Schoenung, S.M., W.V. Hassenzahl. 2003. Long-vs. short-term energy storage technologies analysis. A life-cycle cost study. A study for the DOE energy storage systems program. Sandia National Laboratories.

Shen, Z.J.M., C. Coullard, M.S. Daskin. 2003. A joint location-inventory model. *Transportation Science* **37**(1) 40–55.

Talinli, I., E. Topuz, E. Aydin, S. Kabakci. 2011. A holistic approach for wind farm site selection by using FAHP. [Online; accessed on 17-September-2013].

Taylor, J.A., F.S. Hover. 2013. Conic AC transmission system planning. *Power Systems, IEEE Transactions on* **28**(2) 952–959.

Thomann, G.C., M.J. Barfield. 1988. The time variation of wind speeds and wind farm power output in kansas. *Energy Conversion, IEEE Transactions on* **3**(1) 44–49.

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

E-Companion

EC.1. Analysis of Model Inaccuracy in Subsection 4.4

We have shown in the main text that the modeling process described in Subsection 4.4 incorporates major wind characteristics and leads to efficient ES-transmission planning problem formulation (13). However, these merits come at the cost of three sources of model inaccuracy: (i) The use of uniform distribution approximation of wind outputs at individual wind farms affects the sizing of the transmission lines that are upstream of the junction sites. (ii) The similar approximation applied to the aggregated wind outputs at the junction sites affects the sizing of the downstream transmission lines. (iii) Myopically sizing the upstream transmission lines without considering additional capacity cost downstream of the junction sites creates model suboptimality. The first source of model inaccuracy has been demonstrated in Subsection 4.3 to be reasonably small. In this section, we quantify and explain the second and the third components of model inaccuracy using numerical and/or theoretical analysis. The objectives of this analysis are to show that the model inaccuracy is reasonably small in most practical settings and to gain more insights into the nature of this infrastructure planning problem.

EC.1.1. Errors of Approximating Aggregated Curtailed Wind Output at a Junction Site

We first investigate the approximation errors of the aggregated wind outputs and show that the errors are reasonably small under practical parameter settings. The numerical example in Subsection 4.3 has demonstrated that approximating wind output by uniform distribution incurs reasonably small cost error in the single-farm scenario. In the multiple-farm scenario, however, even if uniformly distributed, the wind outputs at the individual farms get first curtailed by the ES-free transmission lines and then aggregated, resulting in non-uniform wind outputs faced by the junction sites. In order to characterize those aggregated wind outputs as closed-form expressions (12a) and (12b), we applied uniform distribution approximation again - first to the individually curtailed and then to the aggregated outputs.

To investigate this approximation error, we consider the following scenario: Three wind farms generate power flows that are correlated, have the same probability distribution and travel the same distance before getting aggregated at a junction site. This simplified setting is adequate for us to observe the effect of energy curtailment and aggregation. In the meantime, it minimizes the number of parameters to perturb. Given wind power data, each instance of the numerical test can be fully characterized by three parameters: the downstream transmission distance, the correlation coefficient between the wind outputs and the single value of upstream curtailment factor η as defined in (9), which has one-one correspondence to the upstream transmission distance.

Figure EC.1 (a) shows the distribution of the real wind output at a single farm. The spike of the high power results from the rated power limit of wind turbines. The outputs at the other two farms are generated by re-sampling from the same data to have the same probability distribution and the specified correlation. Figures EC.1 (b)-(f) are the histograms of the distributions of the aggregated wind outputs (in dark bars) and their uniform distribution approximations (in light bars) based on (12a) and (12b). We observe that, as expected, either decreasing the inter-farm correlation or increasing the upstream curtailment results in more non-uniform profile of the aggregated output distribution. Notice that the uniform distribution may poorly approximate the real profile if the upstream curtailment is extremely high, as represented by Figure EC.1 (f).

We next define metrics of the approximation error. We first evaluate the downstream variable cost, which is the sum of the transmission capacity cost and the friction (curtailment) cost that are downstream of the ES-coupled (ES-free) junction site, in three different ways: (i) The real cost is generated with real wind output data and with downstream transmission capacity prescribed by (4) ((9)). (ii) The actual optimal cost is generated with the same real data and with the downstream transmission capacity found by line-search that leads to the minimum downstream variable cost. (iii) The approximated cost is given by equation (10) ((5)). Then we use the following two error metrics: (i) The real error refers to the relative error between the real cost and the actual optimal cost; (ii) and the model error refers to the relative error between the approximated cost and the

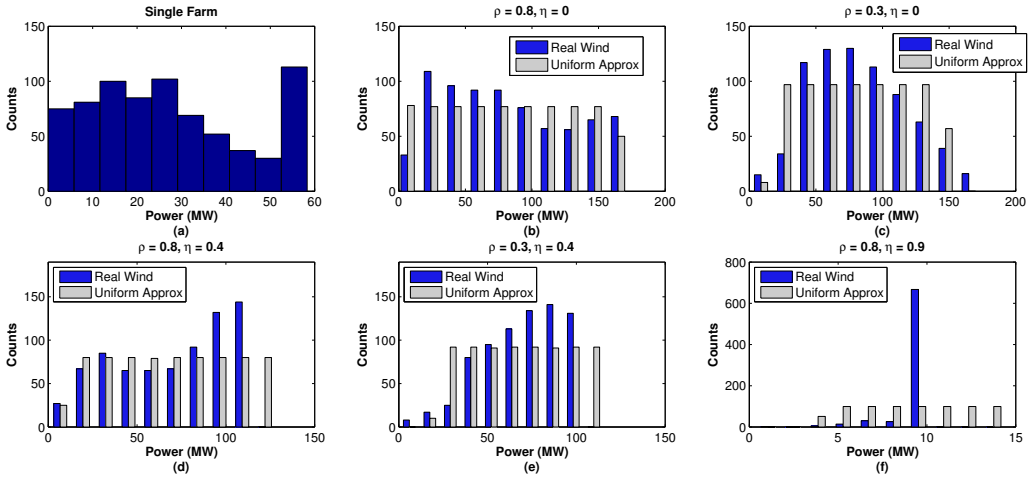

Figure EC.1 Aggregated wind outputs with different degrees of correlation and curtailment.

actual optimal cost. These two metrics, from the cost perspective, are consistent with most of error analysis in the main text. They also isolate the approximation error in the sense that the inaccuracy due to myopically sizing the upstream transmission lines becomes irrelevant, since the upstream transmission capacity is fixed in each instance as given by (9).

For the case where the junction site is co-located with ES, Figure EC.2 shows the cost errors of 300 numerical test instances with different parameter values. One can see that the real cost errors are less than 3% in almost all the instances, which is a strong indication that the prescribed downstream transmission line capacity is near-optimal. When the junction-load distance is long and the degree of the upstream curtailment is high, the real error tends to be zero, since both the prescribed and the actual optimal line capacity tends to be the average wind power. The model errors are in general greater than the real errors, but are still contained within $\pm 10\%$ in most instances.

For the case where the junction site has no ES, Figure EC.3 shows the cost errors of another 300 numerical test instances, from which we make the following observations: (i) Overall, the cost errors are greater than their counterparts in the ES-coupled case. This is because the absence of ES leads to the choice of higher downstream transmission capacity. The uniform distribution approximation is less effective when the line capacity is closer to the maximum wind power, since

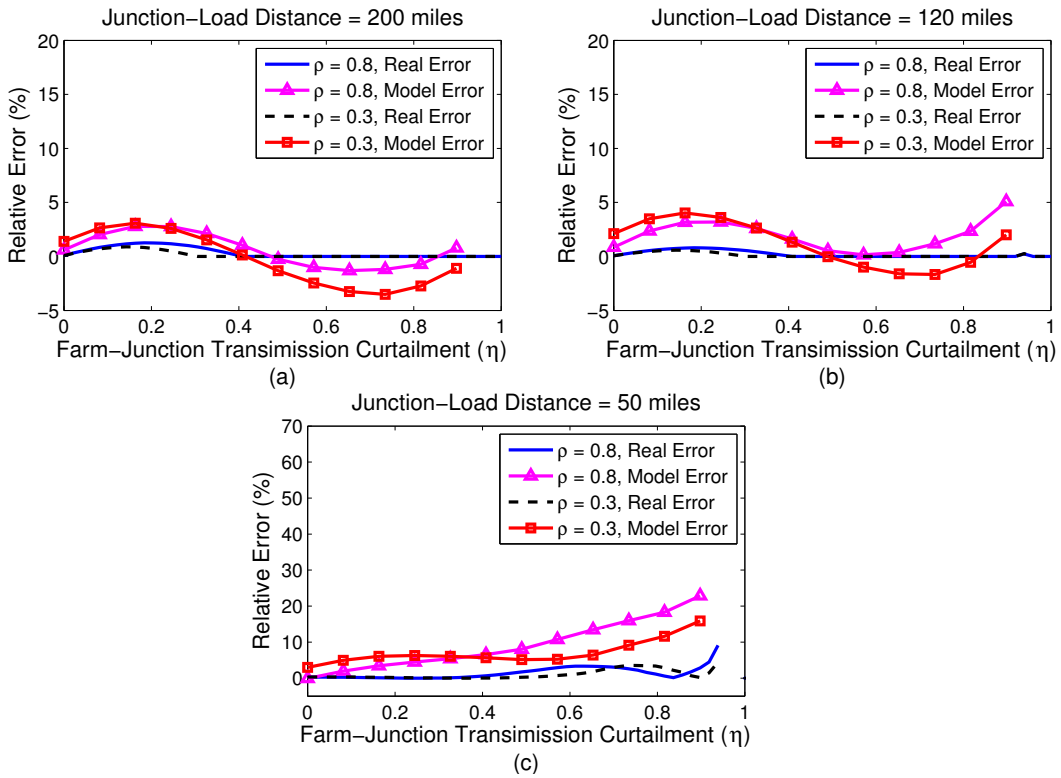

Figure EC.2 Real and model errors of uniform distribution approximation with an ES-coupled junction site.

the curtailment loss, as represented by the right tail of the real wind output distribution, becomes more non-uniform and thus more difficult to approximate. (ii) Instances with lower correlation between individual wind outputs have smaller real and model errors, as a result of the evener output profile as discussed previously. It suggests that spatial pooling enhances the accuracy of uniform distribution approximation by reducing the wind power variability. (iii) The approximation errors (particularly the real errors) are still moderate when the upstream curtailment $\eta < 0.5$. Although we consider the full range of values of η for demonstration purposes, in reality, however, farm-junction transmission distance should be relatively short to save the ES-free transmission line cost and prevent substantial curtailment loss. For example, η is only 0.07 when the farm-junction distance is 50 miles in our numerical test setting; the associated approximation error is very small.

With the above numerical studies, we conclude that the error of approximating wind output distribution using uniform distribution can be reasonably small in practical settings. Spatial pooling of wind outputs and the deployment of ES systems both help enhance this approximation accuracy.

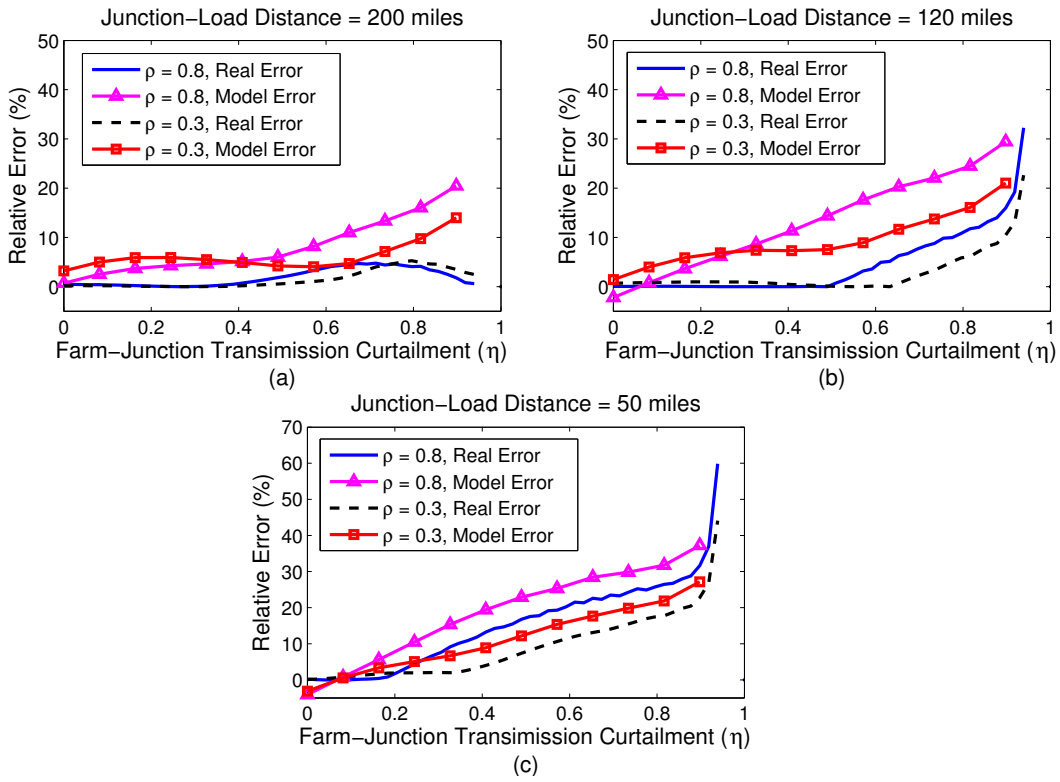

Figure EC.3 Real and model errors of uniform distribution approximation with an ES-free junction site.

EC.1.2. Sub-optimality Due to Myopically Sizing a Transmission Line Upstream of a Junction Site

Another source of model inaccuracy stems from the oversizing of the transmission lines. This subsection describes an analytical approach to quantifying this inaccuracy, and then uses numerical experiments to justify that the inaccuracy is indeed minor.

For tractability reason, planning model (13) incorporates (9) developed in the single-farm scenario as the farm-junction transmission line capacities. These capacity quantities do not take into account the additional transmission capacity cost downstream of the junction sites, and thus tend to be larger-than-optimal. Consequently, with less curtailment loss from upstream, the transmission lines downstream of the junction sites expect to deliver more-than-optimal wind power, so their capacities prescribed by the model is no-less-than-optimal.

To analyze this sub-optimality, we use the following problem setting: Uniformly distributed power is generated from a single wind farm, curtailed and transmitted by an ES-free line down to an ES-

coupled (ES-free) junction site that is l_u miles away, and then flows another $(l-l_u)$ miles through an economic (ES-free) line to get to the load center. This simplified setting not only enables us to analytically quantify the sub-optimality, but also excludes the model inaccuracy that is caused by approximating wind outputs at different stages.

The sub-optimality can be measured as the relative error between the expected real total cost and the expected optimal total cost. To compute the expected real total cost, the capacity of the transmission line upstream of the junction site is given by (9), while that capacity downstream the junction site is the minimizer of the expected total downstream curtailment (friction) and capacity cost. On the other hand, the expected optimal total cost is found by jointly optimizing the upstream and the downstream transmission capacities.

The ES-free Case

For the scenario where the junction site has no ES, the following proposition gives an upper bound of the sub-optimality.

PROPOSITION EC.1. *In the above problem setting with a single wind farm generating power* $\mathbf{w_t} \sim$ *Unif*$(\mu-\frac{\epsilon}{2},\mu+\frac{\epsilon}{2})$*, suppose* $\mu-\frac{\epsilon}{2}\geq 0$ *and there is no ES system at the junction site. Then an upper bound of the sub-optimality due to myopically sizing the upstream transmission line is given by*

(i) $\frac{1}{3}\frac{\eta}{(2-\eta)}$*, if* $l_u\in[0,\frac{l}{2})$*;*

(ii) $\frac{1}{4}\frac{\eta}{(2-\eta)}$*, if* $l_u\in[\frac{l}{2},l]$*;*

where $\eta=\frac{al\delta}{p}<1$*.*

Proof of Proposition EC.1: First consider the scenario of optimally sizing the transmission lines upstream and downstream of the junction site. Obviously, it is never optimal to assign different capacities to these two segments. Therefore, the problem is reduced to the case in Subsection 4.2, where a single wind farm delivers power through a single line to the load center. The expected optimal total cost is given by (10), which we explicitly rewrite here:

$$v^*=al\mu+al(\frac{1}{2}-\frac{1}{2}\frac{\delta al}{p})\epsilon. \tag{EC.1}$$

As to the case of myopically sizing the upstream transmission line, its capacity and the expected upstream cost are given by (9) and (10), respectively, as follows:

$$C_u = \mu + (\frac{1}{2} - \eta_u)\epsilon, \tag{EC.2a}$$

$$v_u = al_u\mu + al_u(\frac{1}{2} - \frac{1}{2}\eta_u)\epsilon, \tag{EC.2b}$$

where $\eta_u = \frac{\delta a l_u}{p}$ indicates the degree of curtailment of the upstream line. Consequently, the junction site faces curtailed wind output $\check{\mathbf{w}}_\mathbf{t}$ that has the following cumulative density function:

$$F_{\check{\mathbf{w}}_\mathbf{t}}(w) = \begin{cases} 0, & \text{if } w < \mu - \epsilon/2; \\ \frac{w-(\mu-\epsilon/2)}{\epsilon}, & \text{if } \mu - \epsilon/2 \le w < C_u; \\ 1, & \text{if } w = C_u. \end{cases} \tag{EC.3}$$

Choosing the quantity of downstream line capacity with uncertain wind power is a newsvendor type of problem. The overage cost is the per-unit capacity cost $\delta a(l - l_u)$ and the underage cost is the per-unit profit $p - \delta a(l - l_u)$. Therefore, the newsvendor critical fractile is $\frac{p-\delta a(l-l_u)}{p}$. Note that $F_{\check{\mathbf{w}}_\mathbf{t}}(\cdot)$ is not left-continuous at $w = C_u = \mu + (\frac{1}{2} - \eta_u)\epsilon$, since $F_{\check{\mathbf{w}}_\mathbf{t}}(C_u^-) = 1 - \eta_u < F_{\check{\mathbf{w}}_\mathbf{t}}(C_u) = 1$. This jump discontinuity leads to the following two cases:

Case (i): $l_u < \frac{1}{2}l$. This is equivalent to the critical fractile $\frac{p-\delta a(l-l_u)}{p} < 1 - \frac{\delta a l_u}{p} = 1 - \eta_u$. Then the downstream transmission capacity is given by:

$$\begin{aligned} C_d &= F^{-1}_{\check{\mathbf{w}}_\mathbf{t}}(\tfrac{p-\delta a(l-l_u)}{p}) \\ &= \mu - \epsilon/2 + \tfrac{p-\delta a(l-l_u)}{p}\epsilon \\ &= \mu + (\tfrac{1}{2} - \tfrac{a(l-l_u)\delta}{p})\epsilon. \end{aligned} \tag{EC.4}$$

The expected downstream curtailment loss is calculated as follows:

$$\begin{aligned} \mathbb{E}[\mathbf{l}_d] &= \textstyle\int_{C_d}^{C_u}(w_t - C_d)\frac{1}{\epsilon}dw_t + \mathbb{P}(\mathbf{w_t} \ge C_u)(C_u - C_d) \\ &= \textstyle\int_{\mu+(\frac{1}{2}-\frac{a l_u\delta}{p})\epsilon}^{\mu+(\frac{1}{2}-\frac{a(l-l_u)\delta}{p})\epsilon}(w_t - \mu - (\frac{1}{2} - \frac{a(l-l_u)\delta}{p})\epsilon)\frac{1}{\epsilon}dw_t + \frac{a l_u\delta}{p}(\mu + (\frac{1}{2} - \frac{a l_u\delta}{p})\epsilon - \mu - (\frac{1}{2} - \frac{a(l-l_u)\delta}{p})\epsilon) \\ &= \tfrac{\epsilon}{2}(\tfrac{a\delta}{p})^2(l^2 - 2l_u l). \end{aligned} \tag{EC.5}$$

Adding the cost terms together yields the expected real total cost:

$$\begin{aligned} v &= v_u + \tfrac{p}{\delta}\mathbb{E}[\mathbf{l}_d] + C_d a(l-l_u) \\ &= al\mu + [\tfrac{al}{2} - \tfrac{a^2\delta}{2p}(3l_u^2 - 2l_u l + l^2)]\epsilon. \end{aligned} \tag{EC.6}$$

Comparing (EC.1) and (EC.6), the absolute sub-optimality is bounded as follows:

$$\begin{aligned} v - v^* &= [\tfrac{al}{2} - \tfrac{a^2\delta}{2p}(3l_u^2 - 2l_u l + l^2)]\epsilon - al(\tfrac{1}{2} - \tfrac{1}{2}\tfrac{\delta al}{p})\epsilon \\ &= \tfrac{\delta a^2}{2p}\epsilon(2l_u l - 3l_u^2) \\ &\le \tfrac{\delta a^2}{6p}\epsilon l^2, \end{aligned} \tag{EC.7}$$

since $(2l_u l - 3l_u^2)$ reaches its maximum, $\frac{1}{3}l^2$, when $l_u = \frac{1}{3}l$. Therefore, the relative suboptimality in case (i) is bounded as follows:

$$\begin{aligned} \tfrac{v-v^*}{v^*} &\le \tfrac{\frac{\delta a^2}{6p}\epsilon l^2}{al\mu + al(\frac{1}{2} - \frac{1}{2}\frac{\delta al}{p})\epsilon} \\ &\le \tfrac{\frac{\delta a^2}{6p}\epsilon l^2}{al\frac{\epsilon}{2} + al(\frac{1}{2} - \frac{1}{2}\frac{\delta al}{p})\epsilon} \quad (\text{since } \mu - \epsilon/2 \ge 0) \\ &= \tfrac{\frac{\delta al}{p}}{6(1 - \frac{\delta al}{2p})} \\ &= \tfrac{1}{3}\tfrac{\eta}{2-\eta}, \end{aligned} \tag{EC.8}$$

where $\eta = \frac{\delta al}{p}$. We thus complete the proof of Proposition EC.1 (i).

Case (ii): $l_u \ge \frac{1}{2}l$. This is equivalent to the critical fractile $\frac{p - \delta a(l-l_u)}{p} \ge 1 - \frac{\delta a l_u}{p} = 1 - \eta_u$. Since $F_{\tilde{\mathbf{w}}_\mathbf{t}}(C_d) < 1 - \eta_u \ \forall C_d < C_u$, we know that the choice of downstream line capacity $C_d = C_u = \mu + (\frac{1}{2} - \eta_u)\epsilon$. Therefore, no curtailment occurs downstream of the junction site. The expected real total cost can thus be calculated as:

$$\begin{aligned} v &= v_u + C_d a(l - l_u) \\ &= al_u\mu + al_u(\tfrac{1}{2} - \tfrac{1}{2}\eta_u)\epsilon + [\mu + (\tfrac{1}{2} - \eta_u)\epsilon]a(l - l_u) \\ &= al\mu + [al(\tfrac{1}{2} - \tfrac{\delta al_u}{p}) + \tfrac{p}{2\delta}(\tfrac{\delta al_u}{p})^2]\epsilon. \end{aligned} \tag{EC.9}$$

Comparing (EC.1) and (EC.9), the absolute sub-optimality is bounded as follows:

$$\begin{aligned} v - v^* &= [al(\tfrac{1}{2} - \tfrac{\delta a l_u}{p}) + \tfrac{p}{2\delta}(\tfrac{\delta a l_u}{p})^2]\epsilon - al(\tfrac{1}{2} - \tfrac{1}{2}\tfrac{\delta a l}{p})\epsilon \\ &= \tfrac{\delta a^2 \epsilon}{2p}(l - l_u)^2 \\ &\leq \tfrac{\delta a^2 \epsilon}{8p} l^2, \end{aligned} \tag{EC.10}$$

where the last inequality is due to the fact that $(v - v^*)$ is monotonically decreasing in l_u for $l_u \in [\frac{1}{2}l, l]$ and is thus maximized when $l_u = \frac{1}{2}l$. Therefore, the relative suboptimality in case (ii) is bounded as follows:

$$\begin{aligned} \tfrac{v - v^*}{v^*} &\leq \frac{\frac{\delta a^2 \epsilon}{8p} l^2}{al\mu + al(\frac{1}{2} - \frac{1}{2}\frac{\delta a l}{p})\epsilon} \\ &\leq \frac{\frac{\delta a^2 \epsilon}{8p} l^2}{al\frac{\epsilon}{2} + al(\frac{1}{2} - \frac{1}{2}\frac{\delta a l}{p})\epsilon} \quad (\text{since } \mu - \epsilon/2 \geq 0) \\ &= \frac{\frac{\delta a l}{p}}{8(1 - \frac{\delta a l}{2p})} \\ &= \tfrac{1}{4}\tfrac{\eta}{2 - \eta}. \end{aligned} \tag{EC.11}$$

We thus complete the proof of Proposition EC.1 (ii). □

The ES-coupled Case

Similar analysis of sub-optimality can be applied to the scenario where the junction site is coupled with an ES system. However, complete analysis turns out to involve excessive case discussion and complicated high-order terms in the bounds expressions, revealing little insight into the problem. Therefore, below we discuss only one case, which is common and is the most representative of the problem, as defined by three conditions: (i) The downstream economic transmission line capacity is greater than the mean of the upstream curtailed wind output. Otherwise the downstream line capacity would have to be the equal to the mean upstream output, which loses the information about how the parameters affect the capacity choice. (ii) The upstream transmission distance accounts for a limited proportion of the entire transmission distance. Otherwise the downstream transmission capacity would tend to be equal to the upstream transmission capacity and the problem would be reduced to the ES-free scenario. (iii) The ES system is not severely inefficient. The conditions and the result of the analysis are formalized in the following proposition:

PROPOSITION EC.2. *In the above problem setting, suppose* $\mu - \frac{\epsilon}{2} \geq 0$ *and there is an ES system at the junction site. Under the conditions (i)* $\mu + \epsilon(\frac{1}{2} - \frac{a(l-l_u)\delta}{p(1-\alpha\beta)}) \geq \mu - \frac{1}{2}\epsilon(\frac{al_u\delta}{p})^2$*, (ii)* $l_u < \frac{l}{2-\alpha\beta}$*, and (iii)* $\alpha\beta \in [0.5, 1)$*, an upper bound of the sub-optimality due to myopically sizing the upstream transmission line is given by* $\frac{1}{9}\frac{\theta}{(2-\theta)}$*, where* $\theta = \frac{al\delta}{(1-\alpha\beta)p}$*.*

Proof of Proposition EC.2: Proving Proposition EC.2 requires the closed-form expressions of the expected real total cost and the expected optimal total cost, which we summarize in the following two lemmas and prove them at the end of this proof:

LEMMA EC.1. *Under conditions (i) and (ii) in Proposition EC.2, the expected real total cost is* $v = al(\mu + \epsilon/2) - \frac{a^2\delta(1-\alpha\beta)l_u^2 + 2\delta a^2(l-l_u)^2}{2p(1-\alpha\beta)}\epsilon$*.*

LEMMA EC.2. *Under conditions (ii) and (iii) in Proposition EC.2, the expected optimal total cost is* $v^* = al(\mu + \epsilon/2) - \frac{a^2\delta[\alpha\beta l(l-2l_u)+l_u^2]}{2p\alpha\beta(1-\alpha\beta)}\epsilon$*.*

Given these two lemmas, the absolute sub-optimality is calculated as follows:

$$\begin{aligned} v - v^* &= al(\mu + \epsilon/2) - \tfrac{a^2\delta(1-\alpha\beta)l_u^2 + 2\delta a^2(l-l_u)^2}{2p(1-\alpha\beta)}\epsilon - al(\mu + \epsilon/2) + \tfrac{a^2\delta[\alpha\beta l(l-2l_u)+l_u^2]}{2p\alpha\beta(1-\alpha\beta)}\epsilon \\ &= \tfrac{a^2\delta\epsilon}{2\alpha\beta(1-\alpha\beta)p}[-\alpha\beta(1-\alpha\beta)l_u^2 - 2\alpha\beta(l-l_u)^2 + \alpha\beta l(l-2l_u) + l_u^2] \\ &= \tfrac{a^2\delta\epsilon}{2\alpha\beta(1-\alpha\beta)p}[-\alpha\beta l^2 + 2\alpha\beta l l_u + (1 - 3\alpha\beta + \alpha^2\beta^2)l_u^2] \\ &= \tfrac{a^2\delta\epsilon}{2\alpha\beta(1-\alpha\beta)p}[l_u^2(1-\alpha\beta)^2 - \alpha\beta(l_u - l)^2], \end{aligned} \tag{EC.12}$$

which is clearly increasing in l_u when $l_u < l$. Following condition (ii), letting $l_u = \frac{l}{2-\alpha\beta}$ generates the following upper bound of the absolute sub-optimality:

$$\begin{aligned} v - v^* &< \tfrac{a^2\delta\epsilon}{2\alpha\beta(1-\alpha\beta)p}[(\tfrac{1-\alpha\beta}{2-\alpha\beta})^2 l^2 - \alpha\beta(\tfrac{\alpha\beta-1}{2-\alpha\beta})^2 l^2] \\ &= \tfrac{a^2\delta\epsilon}{2\alpha\beta p}(\tfrac{1-\alpha\beta}{2-\alpha\beta})^2 l^2. \end{aligned} \tag{EC.13}$$

We next derive a lower bound of v^*. This is equivalent to finding the value of l_u that maximizes $m(l_u) = \alpha\beta l(l-2l_u) + l_u^2$, which is the part of the expression of v^* that involves l_u. Being a quadratic function of l_u, $m(l_u)$ is deceasing for $l_u \in [0, \alpha\beta)$ and increasing for $l_u \in [\alpha\beta, \frac{l}{2-\alpha\beta})$. Following condition (iii) that $\alpha\beta \in [0.5, 1)$, it can be verified that $\alpha\beta > \frac{1}{2}\frac{l}{2-\alpha\beta}$. Therefore, $m(0) = \alpha\beta l^2$ is

greater than $m(\frac{l}{2-\alpha\beta})$ and is thus the maximum of $m(l_u)$ for $l_u < \frac{l}{2-\alpha\beta}$. Subsequently, we obtain the following lower bound of v^*:

$$\begin{aligned} v^* &\geq al(\mu+\epsilon/2) - \tfrac{a^2\delta m(0)}{2p\alpha\beta(1-\alpha\beta)}\epsilon \\ &= al(\mu+\epsilon/2) - \tfrac{a^2\delta l^2}{2p(1-\alpha\beta)}\epsilon \\ &\geq al\epsilon - \tfrac{a^2\delta l^2}{2p(1-\alpha\beta)}\epsilon, \end{aligned} \tag{EC.14}$$

where the last inequality is due to the assumption that $\mu-\epsilon/2\geq 0$. Combining inequalities (EC.13) and (EC.14), we obtain the following upper bound of the suboptimality:

$$\begin{aligned} \tfrac{v-v^*}{v^*} &< \frac{\frac{a^2\delta\epsilon}{2\alpha\beta p}(\frac{1-\alpha\beta}{2-\alpha\beta})^2 l^2}{al\epsilon - \frac{a^2\delta l^2}{2p(1-\alpha\beta)}\epsilon} \\ &= \frac{(\frac{1-\alpha\beta}{2-\alpha\beta})^2}{2\alpha\beta}\frac{\frac{a\delta l}{p}}{1-\frac{a\delta l}{2(1-\alpha\beta)p}} \\ &= \frac{(\frac{1-\alpha\beta}{2-\alpha\beta})^2}{2\alpha\beta}\frac{(1-\alpha\beta)\theta}{1-\frac{\theta}{2}}, \end{aligned} \tag{EC.15}$$

where $\theta = \frac{al\delta}{(1-\alpha\beta)p}$. While (EC.15) already presents an upper bound of the suboptimality, we further simplify it to exclude $\alpha\beta$. Using condition (iii) that $\alpha\beta\in[0.5,1)$, we have $\frac{1-\alpha\beta}{2-\alpha\beta} = 1-\frac{1}{2-\alpha\beta} \leq 1-\frac{1}{2-0.5} = \frac{1}{3}$ and $1-\alpha\beta\leq 0.5$. Therefore, the upper bound can be relaxed and simplified as

$$\tfrac{v-v^*}{v^*} < \frac{(\frac{1}{3})^2}{1}\frac{\frac{1}{2}\theta}{1-\frac{\theta}{2}} = \frac{1}{9}\frac{\theta}{(2-\theta)}. \quad \square \tag{EC.16}$$

Proof of Lemma EC.1: To derive the expected real total cost, we first immediately know from (9) that the upstream transmission capacity is

$$C_u = \mu + (\frac{1}{2} - \frac{al_u\delta}{p})\epsilon, \tag{EC.17}$$

and from (10) that the expected real cost upstream of the junction is

$$v_u = al_u\mu + al_u(\frac{1}{2} - \frac{1}{2}\frac{al_u\delta}{p})\epsilon. \tag{EC.18}$$

Given C_u and the consequent wind output that is curtailed at C_u and uniformly distributed in $[\mu-\frac{1}{2}\epsilon, C_u)$, the expected downstream friction loss as a function of the downstream transmission capacity C_d becomes

$$\begin{aligned}\mathbb{E}[\mathbf{l}_d] &= \textstyle\int_{C_d}^{C_u}(w_t - C_d)(1-\alpha\beta)\frac{1}{\epsilon}dw_t + \mathbb{P}(\mathbf{w_t} \geq C_u)(C_u - C_d)(1-\alpha\beta) \\ &= \textstyle\int_{C_d}^{\mu+(\frac{1}{2}-\frac{al_u\delta}{p})\epsilon}(w_t - C_d)(1-\alpha\beta)\frac{1}{\epsilon}dw_t + \frac{al_u\delta}{p}(\mu+(\frac{1}{2}-\frac{al_u\delta}{p})\epsilon - C_d)(1-\alpha\beta) \\ &= \tfrac{1-\alpha\beta}{2\epsilon}(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon - C_d)^2 + \tfrac{al_u\delta}{p}(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon - C_d)(1-\alpha\beta).\end{aligned} \tag{EC.19}$$

The expected real downstream cost becomes

$$\begin{aligned}v_d(C_d) &= a(l-l_u)C_d + \tfrac{p}{\delta}\mathbb{E}[\mathbf{l}_d] \\ &= a(l-l_u)C_d + \tfrac{1-\alpha\beta}{2\delta\epsilon}(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon - C_d)^2 + pal_u(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon - C_d)(1-\alpha\beta) \\ &= \tfrac{1-\alpha\beta}{2\delta\epsilon}C_d^2 + [-\tfrac{p(1-\alpha\beta)(\mu+(\frac{1}{2}-\frac{al_u\delta}{p})\epsilon)}{\delta\epsilon} - pal_u(1-\alpha\beta) + a(l-l_u)]C_d \\ &+ \tfrac{1-\alpha\beta}{2\delta\epsilon}(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon)^2 + pal_u(\mu+(\tfrac{1}{2}-\tfrac{al_u\delta}{p})\epsilon)(1-\alpha\beta).\end{aligned} \tag{EC.20}$$

Applying the first-order condition, we obtain the minimizer of $v_d(C_d)$

$$C_d^* = \mu + \epsilon(\frac{1}{2} - \frac{a\delta(l-l_u)}{p(1-\alpha\beta)}). \tag{EC.21}$$

Condition (i) imposes that this quantity is greater than the mean of the upstream wind output (which is given by (EC.31)); condition (ii) is equivalent to $\mu+\epsilon(\frac{1}{2}-\frac{a\delta(l-l_u)}{p(1-\alpha\beta)}) < \mu+(\frac{1}{2}-\frac{al_u\delta}{p})\epsilon$, or $C_d^* < C_u$, so that ES is necessary. Therefore, C_d^* is the downstream transmission capacity. Substituting C_d with C_d^* in (EC.20) yields the expected real downstream cost:

$$\begin{aligned}v_d^* &= a(l-l_u)(\mu+\tfrac{1}{2}\epsilon - \tfrac{\delta\epsilon a(l-l_u)}{p(1-\alpha\beta)}) + pal_u[\tfrac{\delta\epsilon a}{p}(\tfrac{l-l_u}{1-\alpha\beta} - l_u)](1-\alpha\beta) \\ &+ \tfrac{p(1-\alpha\beta)}{2\delta\epsilon}[\tfrac{\delta\epsilon a}{p}(\tfrac{l-l_u}{1-\alpha\beta} - l_u)]^2.\end{aligned} \tag{EC.22}$$

Then the expected real total cost is the sum of the expected upstream and downstream real costs as given in (EC.18) and (EC.22):

$$\begin{aligned}v &= v_u + v_d^* \\ &= al(\mu+\tfrac{\epsilon}{2}) - \tfrac{a^2\delta(1-\alpha\beta)l_u^2 + 2\delta a^2(l-l_u)^2}{2p(1-\alpha\beta)}\epsilon. \quad \square\end{aligned} \tag{EC.23}$$

Proof of Lemma EC.2: The evaluation of the expected real total cost needs joint optimization of transmission capacities both upstream and downstream of the junction site, denoted by C_u and C_d, respectively. From (8) we know that the expected upstream cost is

$$v_u(C_u) = \frac{p}{2\epsilon\delta}(\mu + \frac{\epsilon}{2} - C_u)^2 + al_uC_u. \tag{EC.24}$$

Given C_u, C_d and the wind output that is curtailed at C_u and uniformly distributed in $[\mu - \frac{1}{2}\epsilon, C_u)$, the expected downstream friction loss becomes

$$\begin{aligned}\mathbb{E}[\mathbf{l}_d] &= \textstyle\int_{C_d}^{C_u}(w_t - C_d)(1-\alpha\beta)\frac{1}{\epsilon}dw_t + \mathbb{P}(\mathbf{w_t} \geq C_u)(C_u - C_d)(1-\alpha\beta)\\ &= \tfrac{1-\alpha\beta}{2\epsilon}(C_u - C_d)^2 + \tfrac{\mu+\epsilon/2-C_u}{\epsilon}(C_u - C_d)(1-\alpha\beta)\\ &= -\tfrac{1-\alpha\beta}{2\epsilon}C_u^2 + \tfrac{1-\alpha\beta}{2\epsilon}C_d^2 + \tfrac{(\mu+\epsilon/2)(1-\alpha\beta)}{\epsilon}(C_u - C_d).\end{aligned} \tag{EC.25}$$

The expected downstream cost is

$$\begin{aligned}v_d(C_u, C_d) &= a(l-l_u)C_d + \tfrac{p}{\delta}\mathbb{E}[\mathbf{l}_d]\\ &= -\tfrac{p(1-\alpha\beta)}{2\epsilon\delta}C_u^2 + \tfrac{p(\mu+\epsilon/2)(1-\alpha\beta)}{\epsilon\delta}C_u + \tfrac{p(1-\alpha\beta)}{2\epsilon\delta}C_d^2 + [a(l-l_u) - \tfrac{p(\mu+\epsilon/2)(1-\alpha\beta)}{\epsilon\delta}]C_d.\end{aligned} \tag{EC.26}$$

Then the expected total cost becomes

$$\begin{aligned}v(C_u, C_d) &= v_u(C_u) + v_d(C_u, C_d)\\ &= \tfrac{\alpha\beta p}{2\delta\epsilon}C_u^2 + \tfrac{(2a\delta\epsilon l_u - 2\alpha\beta\mu p - \alpha\beta p\epsilon)}{2\delta\epsilon}C_u\\ &+ \tfrac{p(1-\alpha\beta)}{2\epsilon\delta}C_d^2 + [a(l-l_u) - \tfrac{p(\mu+\epsilon/2)(1-\alpha\beta)}{\epsilon\delta}]C_d + \tfrac{p}{2\epsilon\delta}(\mu + \tfrac{\epsilon}{2})^2. \quad \square\end{aligned} \tag{EC.27}$$

Note that the bilinear term C_uC_d cancels out in (EC.25). As a result, the optimal upstream and downstream transmission capacities, C_u^* and C_d^*, do not depend on each other. Applying the first-order condition to (EC.27) with respect to C_u and C_d, respectively, we obtain

$$C_u^* = \mu + \epsilon(\frac{1}{2} - \frac{a\delta l_u}{\alpha\beta p}), \tag{EC.28a}$$

$$C_d^* = \mu + \epsilon(\frac{1}{2} - \frac{a\delta(l-l_u)}{(1-\alpha\beta)p}). \tag{EC.28b}$$

Plugging (EC.28a) and (EC.28b) into (EC.27) yields the expected optimal total cost given as follows:

$$v^* = v(C_u^*, C_d^*) = al(\mu + \epsilon/2) - \frac{a^2\delta[\alpha\beta l(l-2l_u)+l_u^2]}{2p\alpha\beta(1-\alpha\beta)}\epsilon. \quad \square \tag{EC.29}$$

Remarks

In most of practical settings of parameter values, η and θ are both much less than one, suggesting that the model suboptimality as bounded in Propositions (EC.1) and (EC.2) are well-contained. Moreover, these propositions as well as their proofs lead to the following observations about the properties of the problem:

Curtailment-independence of downstream transmission capacity. After myopically sizing the upstream transmission lines in both the ES-free and ES-coupled scenarios, we may choose the downstream transmission capacity as if there were no upstream curtailment at all (though it should not exceed the upstream line capacity). To see this point, notice that, under proper conditions, the downstream transmission capacities prescribed in (EC.4) and (EC.21) are identical with those prescribed by (9) and (4), respectively, where there assumes to be no upstream curtailment. This effect of curtailment-independence might be against the intuition that higher degree of upstream curtailment would result in more conservative sizing of the downstream line. However, it can be understood by noticing that the power that has been curtailed by the upstream line, no matter how much it is, is forgone and irrelevant to the downstream power transmission. In other words, upstream curtailment does not affect the profile of the wind power that can be transmitted via the downstream line.

This property may help simplify planning model (13). Specifically, we may replace those farm-junction-specific wind characteristics (e.g., moments given by (11)) simply with wind characteristics at each wind farm. This simplification frees us from approximating curtailed wind power distribution by uniform distribution, which may also help reduce the approximation error.

Decoupling of the joint optimization of upstream and downstream transmission capacities. Proof of Lemma EC.2 shows that, under the given conditions, the jointly optimal upstream and

downstream transmission capacities given by (EC.28) can actually be obtained in a decoupled fashion, since the cross term of these two quantities cancels out in the expression of the expected total cost (EC.27). The explanation of the curtailment-independence of downstream transmission capacity also applies here.

More interestingly, two consequences of this decoupling effect presents more insights into our planning model. First, the jointly optimal downstream transmission capacity given by (EC.28b) turns out to be identical to that capacity given by (EC.21) following the sub-optimal approach, which is employed in our planning problem formulation. It suggests that, the downstream transmission capacity, which accounts for a major part of the total cost in reality, can be prescribed to be near-optimal by our planning model.

On the other hand, the jointly optimal upstream transmission capacity given by (EC.28a) differs from the myopically determined quantity (EC.17) in the degree of curtailment by a factor of $\frac{1}{\alpha\beta}$. In this way, the optimal sizing of upstream transmission takes the ES conversion efficiency into account, tending to curtail more power if it anticipates more friction loss downstream. Therefore, we may partly offset the suboptimality of planning model (13) by replacing $\eta = \frac{\delta q}{p}$ with $\tilde{\eta} = \frac{\delta q}{\alpha\beta p}$ in the terms of upstream variable cost of the objective function (13a). More specifically, the terms in the second bracket in (13a) can be replaced with the following terms:

$$\sum_{j\in J}\sum_{i\in I}[q_{ij}\mu_i + q_{ij}(\frac{1}{2} - \frac{1}{2}\tilde{\eta}_{ij})\epsilon_i]Y_{ij} + \sum_{j\in J}\sum_{i\in I}[q_{ij}\mu_i + q_{ij}(\frac{1}{2} - \frac{1}{2}\eta_{ij})\epsilon_i]Z_{ij}. \tag{EC.30}$$

Numerical Tests

We next carry out numerical tests for further validation. Note that the optimal total cost is found by doing two-dimensional line-search over the upstream and downstream transmission capacities given $2,000$ hours of uniformly distributed wind power from the wind farm.

In the ES-free scenario, we generate $10,000$ test instances for different values of $l \in [10, 460]$ and $l_u \in [0, l]$. In the ES-coupled scenario, we generate $100,000$ test instances for different values of $l \in [10, 460]$, $l_u \in [0, l]$ and $\alpha\beta \in \{0.3, 0.8\}$. The resulting suboptimality as the relative error of the

real total cost against the optimal total cost is summarized in Table EC.1. Note that the maximum suboptimality in the ES-free scenario, 33.6%, is close to the upper bound $\frac{1}{3}$ given by Proposition (EC.1) when $\eta = \frac{\delta a l}{p} = 1$. When only considering more practical cases where $\eta \leq 0.5$ (i.e., $l < 230$ miles), this suboptimality becomes much smaller in both cases.

The resulting suboptimality averages 2.48% with maximum 25.3%. Again, when only considering those instances with $\eta \leq 0.5$, the suboptimality becomes much smaller, averaging 0.41% with maximum 4.97%.

Table EC.1 Model suboptimality due to myopically sizing an upstream transmission line.

	ES-Free		**ES-Coupled**	
	Mean	Maximum	Mean	Maximum
$\eta \in (0,1)$	5.68%	33.6%	1.34%	5.18%
$\eta \in (0,0.5]$	2.48%	25.3%	0.41%	4.97%

EC.1.3. Overall Model Inaccuracy

The overall model inaccuracy consists of the model suboptimality and the distribution approximation errors of the wind outputs at individual farms and at the junction sites. To numerically investigate the overall model inaccuracy, we use the same three-farm scenario as described in Subsection EC.1.1 that is used to investigate the approximation error of aggregated wind outputs. However, for the purpose of incorporating model suboptimality, we also evaluate the upstream transmission capacity, both myopically using formula (9) and optimally using line-search. The overall model inaccuracy is then measured as the relative error between the real total cost and the optimal total cost.

In the ES-free scenario, we generate $5,000$ instances for different values of $l \in [20, 460]$, $l_u \in [0, l]$ and $\rho \in \{0.3, 0.8\}$. In the ES-coupled scenario, we generate $200,000$ instances for different values of $l \in [20, 460]$, $l_u \in [0, l]$, $\rho \in \{0.3, 0.8\}$ and $\alpha\beta \in [0.5, 0.99]$. The resulting overall model inaccuracy is summarized in Table EC.2. Through the numerical experiments, we observe, again, that large

errors occur in the less realistic scenario where the whole transmission distance is too long to be profitable. For those instances with $\eta \leq 0.5$ in both the ES-free and ES-coupled scenarios, the maximum error is significantly smaller as shown in Table EC.2.

Table EC.2 Overall model inaccuracy.

	ES-Free		**ES-Coupled**	
	Mean	Maximum	Mean	Maximum
$\eta \in (0,1)$	10.7%	39.4%	7.08%	30.7%
$\eta \in (0,0.5]$	8.25%	17.6%	6.42%	14.8%

In summary, we have demonstrated that the model inaccuracy due to uniform distribution approximation of wind outputs and myopically sizing upstream transmission lines is reasonably small in most practical settings. Under some less realistic conditions, however, the model inaccuracy can be substantial. We also provide closed-form upper bounds of model suboptimality and present more insights that may potentially improve our planning model.

EC.2. Proofs

EC.2.1. Derivation of Equations (11a) and (11b)

In Section 5 we use 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 $\eta_{ij} < 1$, we claim that this uniform distribution has mean and interval length as follows:

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

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

To see (11a), note that the expected wind power that is curtailed by a transmission line of capacity C can be expressed as:

$$\begin{aligned}\mu_c &= \textstyle\int_{\mu-\frac{\epsilon}{2}}^{C} w\frac{1}{\epsilon}\mathrm{d}w + C\mathbb{P}(\mathbf{w_t} \geq C) \\ &= \tfrac{1}{\epsilon}\tfrac{1}{2}[C^2 - (\mu-\tfrac{\epsilon}{2})^2] + \tfrac{C}{\epsilon}(\mu+\tfrac{\epsilon}{2}-C) \\ &= C - \tfrac{(C-\mu+\frac{\epsilon}{2})^2}{2\epsilon}.\end{aligned} \tag{EC.32}$$

Substituting C with the optimal quantity $\mu + (\frac{1}{2} - \eta)\epsilon$ as given by (9) yields the expression of the expected curtailed wind power given by (11a).

As for (11b), the second moment of the wind power curtailed by a line with capacity C is:

$$\begin{aligned}\mathbb{E}[\mathbf{w}_{t,c}^2] &= \textstyle\int_{\mu-\frac{\epsilon}{2}}^{C} w^2\frac{1}{\epsilon}\mathrm{d}w + C^2\mathbb{P}(\mathbf{w_t} \geq C) \\ &= \tfrac{1}{\epsilon}\tfrac{1}{3}[C^3 - (\mu-\tfrac{\epsilon}{2})^3] + \tfrac{C^2}{\epsilon}(\mu+\tfrac{\epsilon}{2}-C).\end{aligned} \tag{EC.33}$$

The variance of this curtailed wind power becomes:

$$\begin{aligned}Var(\mathbf{w}_{t,c}) &= \mathbb{E}[\mathbf{w}_{t,c}^2] - \mu_c^2 \\ &= \tfrac{(2C-2\mu+\epsilon)^3(-6C+6\mu+5\epsilon)}{192\epsilon^2}.\end{aligned} \tag{EC.34}$$

Again substituting C with the optimal quantity $\mu + (\frac{1}{2} - \eta)\epsilon$ into the above equation yields:

$$Var(\mathbf{w}_{t,c}) = \frac{(1-\eta)^3(1+3\eta)\epsilon^2}{12}. \tag{EC.35}$$

The interval length of this uniform distribution becomes:

$$\epsilon_c = \sqrt{12Var(\mathbf{w}_{t,c})} = \sqrt{(1-\eta)^3(1+3\eta)}\epsilon, \tag{EC.36}$$

which proves (11b). □

EC.2.2. Proof of Proposition 1

Proposition 1. *Assume $C\delta_b \geq \mathbb{E}[\mathbf{w}_{b,\tau}]$, 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) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) \geq \mathbb{P}(\mathbf{s}_\tau = S)$;

(ii) $\tilde{\mathbb{E}}[\mathbf{o}_\tau] \geq \mathbb{E}[\mathbf{o}_\tau]$,

where $\tilde{\mathbb{P}}(\cdot)$ and $\tilde{\mathbb{E}}[\cdot]$ denote probability and expectation with respect to $\tilde{f}_s$, respectively.

Proof of Proposition 1(i). Throughout the proofs of Propositions 1, 2 and 3, we use $\mathbb{P}_w(\cdot)$ and $\mathbb{E}_w[\cdot]$ to denote probability and expectation with wind output $\mathbf{w_{b,\tau+1}}$ being the underlying random variable, respectively.

In order to prove Proposition 1(i), we first express $\mathbb{P}(\mathbf{s}_\tau = S)$ and $\mathbb{P}(\mathbf{s}_\tau = 0)$ in terms of $f_s(s)$. From (14), which describes the transition of $\mathbf{s}_\tau$, the recursive formulae of $\mathbb{P}(\mathbf{s}_\tau = S)$ and $\mathbb{P}(\mathbf{s}_\tau = 0)$ in the long run are given as follows:

$$\begin{aligned}\mathbb{P}(\mathbf{s}_\tau = S) = &\int_{0^+}^{S^-} \mathbb{P}_w(s + \alpha(\mathbf{w_{b,\tau+1}} - C\delta_b) \geq S) f_s(s)\mathrm{d}s \\ &+\mathbb{P}_w(\mathbf{w_{b,\tau+1}} \geq C\delta_b)\mathbb{P}(\mathbf{s}_\tau = S) + \mathbb{P}_w(\alpha(\mathbf{w_{b,\tau+1}} - C\delta_b) \geq S)\mathbb{P}(\mathbf{s}_\tau = 0),\end{aligned} \tag{EC.37}$$

$$\begin{aligned}\mathbb{P}(\mathbf{s}_\tau = 0) = &\int_{0^+}^{S^-} \mathbb{P}_w(\beta s \leq C\delta_b - \mathbf{w_{b,\tau+1}}) f_s(s)\mathrm{d}s \\ &+\mathbb{P}_w(\mathbf{w_{b,\tau+1}} \leq C\delta_b)\mathbb{P}(\mathbf{s}_\tau = 0) + \mathbb{P}_w(\beta S \leq C\delta_b - \mathbf{w_{b,\tau+1}})\mathbb{P}(\mathbf{s}_\tau = S).\end{aligned} \tag{EC.38}$$

The last term on the RHS of (EC.37) and the last term on the RHS of (EC.38) are zero, since $\mathbb{P}_w(\alpha(\mathbf{w_{b,\tau+1}} - C\delta_b) \geq S) = \mathbb{P}_w(\beta S \leq C\delta_b - \mathbf{w_{b,\tau+1}}) = 0$, according to the first assumption that we make prior to Proposition 1. Rearranging the terms in the above equations yields:

$$\mathbb{P}(\mathbf{s}_\tau = S) = \frac{1}{1 - \mathbb{P}_w(\mathbf{w_{b,\tau+1}} \geq C\delta_b)} \int_{0^+}^{S^-} \mathbb{P}_w(s + \alpha(\mathbf{w_{b,\tau+1}} - C\delta_b) \geq S) f_s(s)\mathrm{d}s, \tag{EC.39}$$

$$\mathbb{P}(\mathbf{s}_\tau = 0) = \frac{1}{1 - \mathbb{P}_w(\mathbf{w_{b,\tau+1}} \leq C\delta_b)} \int_{0^+}^{S^-} \mathbb{P}_w(\beta s \leq C\delta_b - \mathbf{w_{b,\tau+1}}) f_s(s)\mathrm{d}s. \tag{EC.40}$$

For notational simplicity, let

$$A(s) = \frac{\mathbb{P}_w(s + \alpha(\mathbf{w_{b,\tau+1}} - C\delta_b) \geq S)}{1 - \mathbb{P}_w(\mathbf{w_{b,\tau+1}} \geq C\delta_b)}, \tag{EC.41}$$

$$B(s) = 1 + \frac{\mathbb{P}_w(\beta s \leq C\delta_b - \mathbf{w_{b,\tau+1}})}{1 - \mathbb{P}_w(\mathbf{w_{b,\tau+1}} \leq C\delta_b)}. \tag{EC.42}$$

We thus have $\mathbb{P}(\mathbf{s}_\tau = S) = \int_{0^+}^{S^-} A(s) f_s(s)\mathrm{d}s$ and $\mathbb{P}(\mathbf{s}_\tau = 0) = \int_{0^+}^{S^-} (B(s) - 1) f_s(s)\mathrm{d}s$. Notice that, in the interval $(0, S)$, $A(s)$ is always positive and increasing in s, whereas $B(s)$ is always positive

and decreasing in s. Then substituting (EC.39)-(EC.42) into the identity $\mathbb{P}(\mathbf{s}_\tau = S) + \int_{0^+}^{S^-} f_s(s)\mathrm{d}s + \mathbb{P}(\mathbf{s}_\tau = 0) = 1$, we obtain

$$\int_{0^+}^{S^-} (A(s) + B(s)) f_s(s)\mathrm{d}s = 1. \tag{EC.43}$$

We repeat the above procedure from (EC.37) to (EC.42), with $\mathbb{P}(\cdot)$ and $f_s(\cdot)$ being replaced with $\tilde{\mathbb{P}}(\cdot)$ and $\tilde{f}_s(\cdot)$, respectively. Notice that both $A(s)$ and $B(s)$ are functions of probabilities of wind outputs, which are independent from any probability distribution of storage level. Therefore, the resulting expressions of $A(s)$ and $B(s)$ are identical to those given by (EC.41) and (EC.42). Since $\tilde{f}_s(s)$ is also defined as pdf of the storage level, we thus obtain

$$\int_{0^+}^{S^-} (A(s) + B(s)) \tilde{f}_s(s)\mathrm{d}s = 1. \tag{EC.44}$$

Comparing Eqns.(EC.43) and (EC.44) gives

$$\int_{0^+}^{S^-} (A(s) + B(s))[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s = 0. \tag{EC.45}$$

We have assumed that $f_s(s)$ is decreasing in s on $(0, S)$. This assumption can be reasoned as follows: we have chosen the interval length δ_b such that $\{\mathbf{w}_{b,\tau}\}$ is an i.i.d process. And notice that $C\delta_b$ is greater than or equal to the mean of $\mathbf{w}_{b,\tau}$. Therefore, at certain storage level s_τ, the probability of charge is greater than or equal to the probability of discharge. Also notice that $\alpha, \beta \in (0, 1)$, indicating that 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 $C\delta_b < \mathbf{w}_{b,\tau}$ because of the friction loss. We thus know that $f_s(s_\tau + \Delta S|s_\tau) < f_s(s_\tau - \Delta S|s_\tau), \forall s_\tau, \Delta S$ satisfying $0 < s_\tau - \Delta S < s_\tau < s_\tau + \Delta S < S$. We do not establish a similar inequality of conditional probability given $s_\tau = 0$ or S. However, since S is assumed to be large enough and real wind output distribution is smooth and continuous, unconditioned density function $f_s(s)$ is tested to be decreasing in s on the open interval $(0, S)$.

With the above monotonic assumption, since $\tilde{f}_s(s)$ is constant, and $A(s), B(s) > 0$ on $(0, S)$, we further know from (EC.45) that $f_s(s) - \tilde{f}_s(s)$ is decreasing in s and cross zero once at the point denoted by $s_{mid} \in (0, S)$.

To show $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) \geq \mathbb{P}(\mathbf{s}_\tau = S)$, it is equivalent to showing $\int_{0^+}^{S^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \leq 0$. Suppose, in order to derive a contradiction, that $\int_{0^+}^{S^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s > 0$, which also implies that $\int_{0^+}^{S^-} B(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s < 0$ from (EC.45). Construct a constant function $\bar{A}(s) \equiv A(s_{mid})$. Then $\bar{A}(s) > A(s), \forall s \in (0, s_{mid})$ and $\bar{A}(s) < A(s), \forall s \in (s_{mid}, S)$. But we also know that $f_s(s) - \tilde{f}_s(s) > 0, \forall s \in (0, s_{mid})$ and $f_s(s) - \tilde{f}_s(s) < 0, \forall s \in (s_{mid}, S)$. It thus follows that:

$$\begin{aligned}
&\int_{0^+}^{S^-} \bar{A}(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
>\ &\int_{0^+}^{s_{mid}^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s + \int_{s_{mid}}^{S^-} \bar{A}(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
>\ &\int_{0^+}^{s_{mid}^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s + \int_{s_{mid}}^{S^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
=\ &\int_{0^+}^{S^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
>\ &0.
\end{aligned} \tag{EC.46}$$

Since $\bar{A}(s) \equiv A(s_{mid}) > 0$, the above inequality immediately implies that $\int_{0^+}^{S^-} [f_s(s) - \tilde{f}_s(s)]\mathrm{d}s > 0$. However, following the similar approach, we also obtain that $\int_{0^+}^{S^-} [f_s(s) - \tilde{f}_s(s)]\mathrm{d}s < 0$ from the hypothesis $\int_{0^+}^{S^-} B(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s < 0$. Therefore, a contradiction is derived, which finishes the proof that $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) \geq \mathbb{P}(\mathbf{s}_\tau = S)$. $\square$

Proof of Proposition 1(ii). Let Λ denote the event that overflow occurs in the next interval. Then from Eqn.(EC.41) we have $\mathbb{P}_w(s + \alpha(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} - C\delta_b) \geq S) = \mathbb{P}_w(\Lambda|s)$ and $A(s) = \frac{\mathbb{P}_w(\Lambda|s)}{1-\mathbb{P}_w(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}} \geq C\delta_b)}$.

Conditioning on whether the storage level is at or below S, $\mathbb{E}[\mathbf{o}_\tau]$ and $\tilde{\mathbb{E}}[\mathbf{o}_\tau]$ can be expressed as follows:

$$\mathbb{E}[\mathbf{o}_\tau] = \mathbb{E}_w[\mathbf{o}_\tau|S]\mathbb{P}(\mathbf{s}_\tau = S) + \int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s) f_s(s)\mathrm{d}s, \tag{EC.47}$$

$$\tilde{\mathbb{E}}[\mathbf{o}_\tau] = \mathbb{E}_w[\mathbf{o}_\tau|S]\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) + \int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s) \tilde{f}_s(s)\mathrm{d}s. \tag{EC.48}$$

In order to show $\tilde{\mathbb{E}}[\mathbf{o}_\tau] \geq \mathbb{E}[\mathbf{o}_\tau]$ by comparing the right-hand sides of (EC.47) and (EC.48), first notice that $\mathbb{E}_w[\mathbf{o}_\tau|S]\mathbb{P}(\mathbf{s}_\tau = S) \leq \mathbb{E}_w[\mathbf{o}_\tau|S]\tilde{\mathbb{P}}(\mathbf{s}_\tau = S)$. Therefore, it suffices to show $\int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \leq 0$, or, equivalently, $\int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \leq 0$. To see why the latter statement is true, notice that $\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]$ is positive and non-decreasing in s, and $A(s)$ is positive. Therefore, $\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]A(s) \leq \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s), \forall s \in (0, s_{mid})$ and $\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]A(s) \geq \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s), \forall s \in (s_{mid}, S)$. Again we already have $f_s(s) - \tilde{f}_s(s) > 0, \forall s \in (0, s_{mid})$ and $f_s(s) - \tilde{f}_s(s) < 0, \forall s \in (s_{mid}, S)$. It thus follows that:

$$
\begin{aligned}
&\int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
&\leq \int_{0^+}^{s_{mid}^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s + \int_{s_{mid}}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
&\leq \int_{0^+}^{s_{mid}^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s + \int_{s_{mid}}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
&= \int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}]A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
&= \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}] \int_{0^+}^{S^-} A(s)[f_s(s) - \tilde{f}_s(s)]\mathrm{d}s \\
&\leq 0.
\end{aligned}
\tag{EC.49}
$$

where the last inequality is due to $\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s_{mid}] > 0$ and the proof by contradiction in (i). Subsequently, we obtain $\tilde{\mathbb{E}}[\mathbf{o}_\tau] \geq \mathbb{E}[\mathbf{o}_\tau]$. □

EC.2.3. Proof of Proposition 2

PROPOSITION 2. *Assume* $\mathbf{w}_{b,\tau} \sim$ *unif*$(\mu_b - \frac{\epsilon_b}{2}, \mu_b + \frac{\epsilon_b}{2})$. *Let* $A = \frac{(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)}$ *and* $B = \frac{\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2(C\delta_b - \mu_b + \frac{\epsilon_b}{2})}$. *Then*

(i) For $s \in (0, S)$, $\tilde{f}_s(s) \equiv f_s^c = \frac{1}{A+S+B}$;

(ii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = 0) = \frac{A}{A+S+B}$;

(iii) $\tilde{\mathbb{P}}(\mathbf{s}_\tau = S) = \frac{B}{A+S+B}$.

Proof of Proposition 2. Given $\mathbf{w}_{b,\tau} \sim \mathrm{unif}(\mu_b - \frac{\epsilon_b}{2}, \mu_b + \frac{\epsilon_b}{2})$ and $\mathbf{w}_{b,\tau}$ is i.i.d, we assume $C\delta_b < \mu_b + \frac{\epsilon_b}{2}$ and $C\delta_b >= \mu_b$ to be nontrivial. Following the transition model (14), we express the conditional probability distribution of the storage level as follows:

$$\begin{aligned}\mathbb{P}_w\{\mathbf{s}_{\tau+\mathbf{1}}=0|s_\tau\}=&\mathbb{P}_w\{C\delta_b-\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}\geq\beta s_\tau\}\\ =&\mathbb{P}_w\{\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}\leq C\delta_b-\beta s_\tau\}\\ =&\begin{cases}0, & \text{if } C\delta_b-\beta s_\tau<\mu_b-\frac{\epsilon_b}{2};\\ \frac{1}{\epsilon_b}[C\delta_b-\beta s_\tau-(\mu_b-\frac{\epsilon_b}{2})], & \text{if } \mu_b-\frac{\epsilon_b}{2}\leq C\delta_b-\beta s_\tau<\mu_b+\frac{\epsilon_b}{2}.\end{cases}\end{aligned} \tag{EC.50}$$

$$\begin{aligned}\mathbb{P}_w\{\mathbf{s}_{\tau+\mathbf{1}}=S|s_\tau\}=&\mathbb{P}_w\{s_\tau+(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}-C\delta_b)\alpha\geq S\}\\ =&\mathbb{P}_w\{\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}\geq\frac{S-s_\tau}{\alpha}+C\delta_b\}\\ =&\begin{cases}0, & \text{if } \frac{S-s_\tau}{\alpha}+C\delta_b>\mu_b+\frac{\epsilon_b}{2};\\ \frac{1}{\epsilon_b}[\mu_b+\frac{\epsilon_b}{2}-\frac{S-s_\tau}{\alpha}-C\delta_b], & \text{if } \mu_b-\frac{\epsilon_b}{2}\leq\frac{S-s_\tau}{\alpha}+C\delta_b\leq\mu_b+\frac{\epsilon_b}{2}.\end{cases}\end{aligned} \tag{EC.51}$$

To derive $f_{\mathbf{s}_{\tau+\mathbf{1}}|s_\tau}(u|s_\tau)$, for infinitesimal storage level increment δu,

$$\begin{aligned}&\mathbb{P}_w\{\mathbf{s}_{\tau+\mathbf{1}}\in[u,u+\delta u]|s_\tau\}\\ =&\begin{cases}\mathbb{P}_w\{s_\tau+\alpha(\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}-C\delta_b)\in[u,u+\delta u]\} & \text{if } u\geq s_\tau\\ \mathbb{P}_w\{s_\tau-\frac{1}{\beta}(C\delta_b-\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}})\in[u,u+\delta u]\} & \text{if } u<s_\tau\end{cases}\\ =&\begin{cases}\mathbb{P}_w\{\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}\in\frac{[u,u+\delta u]-s_\tau}{\alpha}+C\delta_b\}, & \text{if } u\geq s_\tau;\\ \mathbb{P}_w\{\mathbf{w}_{\mathbf{b},\tau+\mathbf{1}}\in([u,u+\delta u]-s_\tau)\beta+C\delta_b\}, & \text{if } u<s_\tau.\end{cases}\end{aligned} \tag{EC.52}$$

Therefore,

$$f_{\mathbf{s}_{\tau+\mathbf{1}}|s_\tau}(u|s_\tau)=\begin{cases}\frac{1}{\alpha\epsilon_b}, & \text{if } u\in[s_\tau,\min\{S,(\mu_b+\frac{\epsilon_b}{2}-C\delta_b)\alpha+s_\tau\});\\ \frac{\beta}{\epsilon_b}, & \text{if } u\in(\max\{0,\frac{\mu_b-\frac{\epsilon_b}{2}-C\delta_b}{\beta}+s_\tau\},s_\tau);\\ 0, & \text{otherwise.}\end{cases} \tag{EC.53}$$

To obtain the approximated unconditional probability distribution of S_τ, notice that $f_{\mathbf{s}_{\tau+\mathbf{1}}}(u)=f_{s_\tau}(u)$ when τ is sufficiently large. Hence, in the case of uniformly distributed wind energy $\mathbf{w}_{b,\tau}$ as well as the constant probability density $\tilde{f}_{s_\tau}(s)\equiv f_s^c,\forall s\in(0,S)$, we have

$$\begin{aligned}
&\tilde{\mathbb{P}}(S_\tau = 0) = \tilde{\mathbb{P}}(\mathbf{s}_{\tau+\mathbf{1}} = 0) \\
&= \int_{0+}^{\frac{C\delta_b - \mu_b + \frac{\epsilon_b}{2}}{\beta}} \frac{C\delta_b - \beta s - \mu_b + \frac{\epsilon_b}{2}}{\epsilon_b} \tilde{f}_{s_\tau}(s)\mathrm{d}s + \tilde{\mathbb{P}}(S_\tau = 0)\frac{C\delta_b - \mu_b + \frac{\epsilon_b}{2}}{\epsilon_b} \\
&= \frac{f_s^c(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta\epsilon_b} + \tilde{\mathbb{P}}(S_\tau = 0)\frac{C\delta_b - \mu_b + \frac{\epsilon_b}{2}}{\epsilon_b}
\end{aligned}$$

$$\Rightarrow \quad \tilde{\mathbb{P}}(S_\tau = 0) = f_s^c \frac{(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta(\mu_b + \epsilon_b/2 - C\delta_b)}, \tag{EC.54}$$

and

$$\begin{aligned}
&\tilde{\mathbb{P}}(S_\tau = S) = \tilde{\mathbb{P}}(\mathbf{s}_{\tau+\mathbf{1}} = S) \\
&= \int_{S-\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)}^{S^-} \frac{\mu_b + \frac{\epsilon_b}{2} - \frac{S-s}{\alpha} - C\delta_b)}{\epsilon_b} \tilde{f}_{s_\tau}(s)\mathrm{d}s + \tilde{\mathbb{P}}(S_\tau = S)\frac{\mu_b + \frac{\epsilon_b}{2} - C\delta_b}{\epsilon_b} \\
&= \frac{f_s^c \alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2\epsilon_b} + \tilde{\mathbb{P}}(S_\tau = S)\frac{\mu_b + \frac{\epsilon_b}{2} - C\delta_b}{\epsilon_b}
\end{aligned}$$

$$\Rightarrow \quad \tilde{\mathbb{P}}(S_\tau = S) = f_s^c \frac{\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2(C\delta_b - \mu_b + \epsilon_b/2)}. \tag{EC.55}$$

Substituting (EC.54) and (EC.55) into the identity $\tilde{\mathbb{P}}(S_\tau = 0) + Sf_s^c + \tilde{\mathbb{P}}(S_\tau = S) = 1$, we obtain

$$f_s^c[\frac{(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta(\mu_b + \epsilon_b/2 - C\delta_b)} + S + \frac{\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2(C\delta_b - \mu_b + \epsilon_b/2)}] = 1.$$

Let $A = \frac{(C\delta_b - \mu_b + \frac{\epsilon_b}{2})^2}{2\beta(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)}$ and $B = \frac{\alpha(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2}{2(C\delta_b - \mu_b + \frac{\epsilon_b}{2})}$.

We thus obtain

$$f_s^c = \frac{1}{A + B + S}, \tag{EC.56}$$

which, together with (EC.54) and (EC.55), finishes the proof of Proposition 2. □

EC.2.4. Proof of Proposition 3

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

Proof of Proposition 3. From Proposition 1(ii) we already know that $\mathbb{E}[\mathbf{o}_\tau]$ is bounded from above by $\tilde{\mathbb{E}}[\mathbf{o}_\tau]$. In the case of uniformly distributed wind energy $\mathbf{w}_{b,\tau}$, we next compute the upper bounds for the terms on the right-hand side of (EC.48), respectively.

For the first term,

$$\mathbb{E}_w[\mathbf{o}_\tau|S] = \int_{C\delta_b}^{\mu_b+\frac{\epsilon_b}{2}} (w - C\delta_b)\frac{1}{\epsilon_b}\mathrm{d}w = \frac{1}{2\epsilon_b}(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2. \tag{EC.57}$$

From Proposition 2,

$$\begin{aligned}
&\tilde{\mathbb{P}}(\mathbf{s}_\tau = S)\\
=\ & f_s^c \tfrac{\alpha(\mu_b+\frac{\epsilon_b}{2}-C\delta_b)^2}{2(C\delta_b-\mu_b+\epsilon_b/2)}\\
<\ & \tfrac{1}{S}\tfrac{\alpha(\mu_b+\frac{\epsilon_b}{2}-C\delta_b)^2}{2(C\delta_b-\mu_b+\epsilon_b/2)}\\
<\ & \tfrac{1}{S}\tfrac{\alpha(\epsilon_b/2)^2}{2\epsilon_b/2} \quad (\text{since } \ C\delta_b \in [\mu_b, \mu_b+\epsilon_b/2])\\
=\ & \tfrac{\alpha\epsilon_b}{4S}.
\end{aligned} \tag{EC.58}$$

For the second term, when $s \in [S - (\mu_b + \epsilon_b/2 - C\delta_b)\alpha, S)$,

$$\begin{aligned}
&\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s) = \mathbb{E}_w[\mathbf{o}_\tau|s]\\
=\ & \textstyle\int_{C\delta_b+(S-s)/\alpha}^{\mu_b+\epsilon_b/2}[w - C\delta_b - (S-s)/\alpha]\frac{1}{\epsilon_b}\mathrm{d}w\\
=\ & \tfrac{[\mu_b+\epsilon_b/2-C\delta_b-(S-s)/\alpha]^2}{2\epsilon_b}.
\end{aligned} \tag{EC.59}$$

Otherwise $\mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s) = 0$. It thus follows that

$$\begin{aligned}
&\textstyle\int_{0^+}^{S^-} \mathbb{E}_w[\mathbf{o}_\tau|\Lambda, s]\mathbb{P}_w(\Lambda|s)\tilde{f}_s(s)\mathrm{d}s\\
=\ & \textstyle\int_{S-(\mu_b+\epsilon_b/2-C\delta_b)\alpha}^{S^-} \frac{[\mu_b+\epsilon_b/2-C\delta_b-(S-s)/\alpha]^2}{2\epsilon_b}\tilde{f}_s(s)\mathrm{d}s\\
<\ & \textstyle\int_{S-(\mu_b+\epsilon_b/2-C\delta_b)\alpha}^{S^-} \frac{[\mu_b+\epsilon_b/2-C\delta_b-(S-s)/\alpha]^2}{2\epsilon_b}\frac{1}{S}\mathrm{d}s\\
=\ & \tfrac{\alpha(\mu_b+\epsilon_b/2-C\delta_b)^3}{6\epsilon_b S}\\
\le\ & \tfrac{\alpha}{6\epsilon_b S}\tfrac{\epsilon_b}{2}(\mu_b+\epsilon_b/2-C\delta_b)^2 \quad (\text{since } \ C\delta_b \ge \mu_b)\\
=\ & \tfrac{\alpha(\mu_b+\epsilon_b/2-C\delta_b)^2}{12S}.
\end{aligned} \tag{EC.60}$$

Substituting (EC.57), (EC.58) and (EC.60) into (EC.48), we obtain

$$\tilde{\mathbb{E}}[\mathbf{o}_\tau] < \frac{5\alpha}{24S}(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2.$$

which provides an upper bound for $\mathbb{E}[\mathbf{o}_\tau]$ as a quadratic function of C. The derivatives of this upper bound are $-\frac{5\alpha}{24S^2}(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2$ with respect to S and $-\frac{5\alpha\delta_b}{12S}(\mu_b + \frac{\epsilon_b}{2} - C\delta_b)^2$ with respect to C. □

EC.2.5. Proof of Proposition 4

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 Subsection 5.3.*

Proof of Proposition 4. To justify the first half of Proposition 4, note that the real optimal ES-transmission deployment is a feasible solution to model (13). Therefore, its total cost v_3 in the uncapacitated scenario given by (13a) is bounded from below by the optimal objective value of model (13). But v_3 is still less than the real optimal cost, since the latter also incorporates the cost of additional ES capacity and the cost of energy overflow. As for the second half, note that the heuristic outlined in Subsection 5.3 generates a feasible deployment but also a total cost that includes overestimated cost of additional ES capacity and the cost of energy overflow. Therefore, this total cost must be greater than the real optimal cost. □

EC.3. Further Inspection of the Upper Bound of ES Capacity in Section 5

In Subsection 4.3 and Section EC.1, we have analyzed the inaccuracy of the model developed in Section 4, where a lower bound of the total cost results from the assumption that the ES systems are uncapacitated. In this section, our goal is to examine the tightness and the accuracy of the cost upper bound as given in Table 5 developed in Section 5, which derives an upper bound for ES capacity. Again we focus on a single wind farm scenario. The cost error analysis for multiple farms is similar and has been presented in the case study in Subsection 6.1.

To set up the experiment, we consider those 24 sites in the case study and look into their wind output data for each of the twelve months in 2006. The cycle length is set to be $\delta_b = 24$hr. The theoretical lower and upper bounds for the investment cost and energy loss are calculated based on the models in Sections 4 and 5, respectively. The economic transmission and the overestimated ES capacities are computed based on Table 5. In order to obtain the actual optimal total cost, we simulate 2000hr system operations for multiple runs, each run with different values of C and S. Then we calculate the average total cost for each run, and select the pair of (C, S) that results in the smallest average total cost. The simulations are carried out with different per-kWh ES capacity costs and with different distances between the wind farm and the load center. Table EC.3 summarizes the average relative 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. Figure EC.4 shows the profiles of site 26784 in March as typical instances. From these results we make the following observations.

Table EC.3 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	23.6%	6.6%	89.7%
120	20.2%	5.17%	81.5%
50	3.09%	6.78%	31.8%

• Firstly, as to the tightness of the cost upper bound, we first observe from Table EC.3 and Figures EC.4 (a), (c) and (e) that the optimal cost in most cases is much closer to the upper bound than to the lower bound, suggesting that the cost upper bound and its associated transmission and ES capacities are more effective estimates of the optimal values. In some cases, the actual optimal cost even exceeds the upper bound, as observed in Figure EC.4 (e). This is due to the inaccuracy of the upper bound which we investigate later in this section. Also notice that, in order to focus

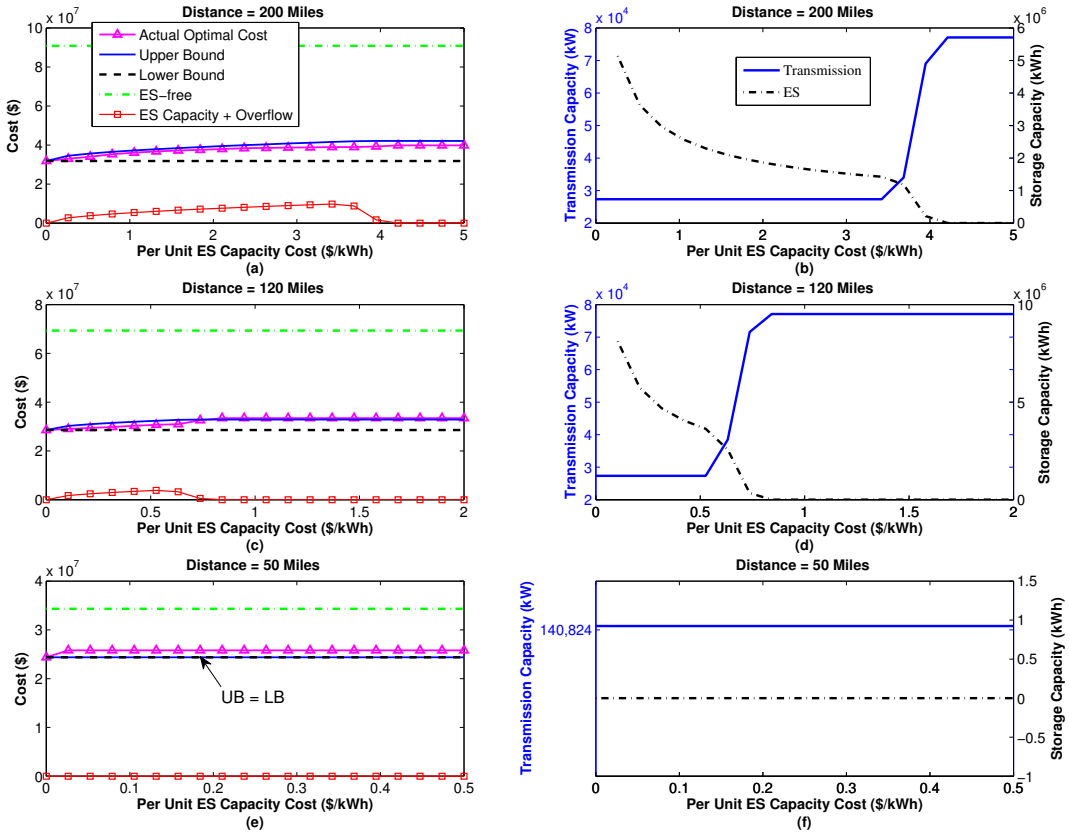

Figure EC.4 Total variable costs ((a), (c), (e)) and economic transmission and (overestimated) ES capacities ((b),(d),(f)) with different distances between a wind farm and a load center.

on the variables of interest, we exclude the fixed transmission and ES installation costs, which, if included, would account for a significant share of the total cost and further narrow the relative gaps. Moreover, the cost gaps narrow as the distance between the wind farm and the load center decreases, because less distance results in cheaper transmission, larger transmission capacity to reduce energy overflow loss and thus less dependence on ES.

- Secondly, the cost gaps are bounded. The cost upper bound increases and then plateaus as the per-kWh ES capacity cost increases, as shown in Figures EC.4 (a), (c) and (e). This is because increase in ES cost makes the investment in transmission capacity more cost-effective than in ES capacity; eventually, the transmission capacity becomes large enough to deliver all the daily wind energy and thus eliminate the use of additional ES capacity for overflow prevention. This ES-

transmission interrelation is clearly demonstrated in Figure EC.4(b). When the ES capacity cost is small ($r \leq \$3.4$/kWh), the transmission capacity C remains at its lowest level of the average wind power μ, since the costly long-distance transmission investment dominates the other cost components and the ES capacity S can be large to hedge against energy overflow. As the ES capacity cost increases, S has to decrease and it becomes more favorable to invest in transmission until C reaches $\frac{\mu_b+\frac{\epsilon_b}{2}}{\delta_b}$ and no additional ES capacity is required when $r \geq \$4.2$/kWh. Hence, as shown in the square-dotted line in Figure EC.4(a), the cost of ES capacity investment and the expected energy overflow first increases and then diminishes to zero.

- Thirdly, the model of cost upper bound involves two sources of inaccuracy. First, the uniform distribution approximation of the daily wind output overlooks the tail of the real output distribution beyond the upper support of the approximated uniform distribution. Therefore, the approximated energy overflow is curtailed. The second source of inaccuracy is the assumption that the wind outputs are independent across days. This assumption is the key to the derivation of the closed-form upper bound expression for ES capacity, which helps deliver managerial insights. However, while choosing long intervals (such as one day) significantly reduces the auto-correlation of the wind output process, ignoring this auto-correlation still causes underestimation of energy overflow loss when storage level is nearly full. Collectively, the underestimation of inter-period overflow loss offsets the overestimation of intra-overflow loss and ES capacity to varying degrees. The resulting cost upper bound is closer to the actual optimal cost than theoretically expected on one hand, but on the other hand, this upper bound can be exceeded.

Figures EC.5 (a)-(c) are the histograms showing the distributions of the lag-one autocorrelation of the daily wind output processes at the 24 sites of twelve months with different transmission distances. The figures also show the proportion of instances where the actual optimal cost exceeds its theoretical upper bound by a certain margin, as represented by the light-colored part of each bin. As expected, the larger the auto-correlation is, the more likely the actual optimal cost is to exceed its upper bound.

We also observe tradeoffs between accuracy and tightness of the cost upper bound. When the transmission distance is as short as 50 miles, Figure EC.5 (c) shows that the upper bounds are exceeded by over 8% in nearly half of all the instances. This is because transmission investment is cheap enough and its capacity is prescribed as $\mu+\rho(\frac{1}{2}-\theta)$ by both the models with and without ES capacity limit. The ES capacity S in those instances also becomes zero. Consequently, the gap between the upper and the lower bounds is zero, but it leaves the underestimation of the overflow dominant and the actual optimal cost above the upper bound. In fact, $S=0$ violates our assumption in Section 5 that S should be large enough such that the probability of the storage level switching from 0 to full state (or the other way round) is negligible. Fortunately, $S=0$ represents the cases where the investment in ES capacity as well as the impact of this inaccuracy is small. If we only consider those instances with $S>0$ due to cheap ES capacity cost, almost no cost upper bounds are violated, as shown in Figure EC.5 (d).

In contrast, Figures EC.5 (a) and (b) and Table EC.3 show that, when the transmission distance $l \geq 120$ miles, the cost upper bounds are not violated in most of the instances, at the cost of relatively large gaps between the upper and the lower bounds.

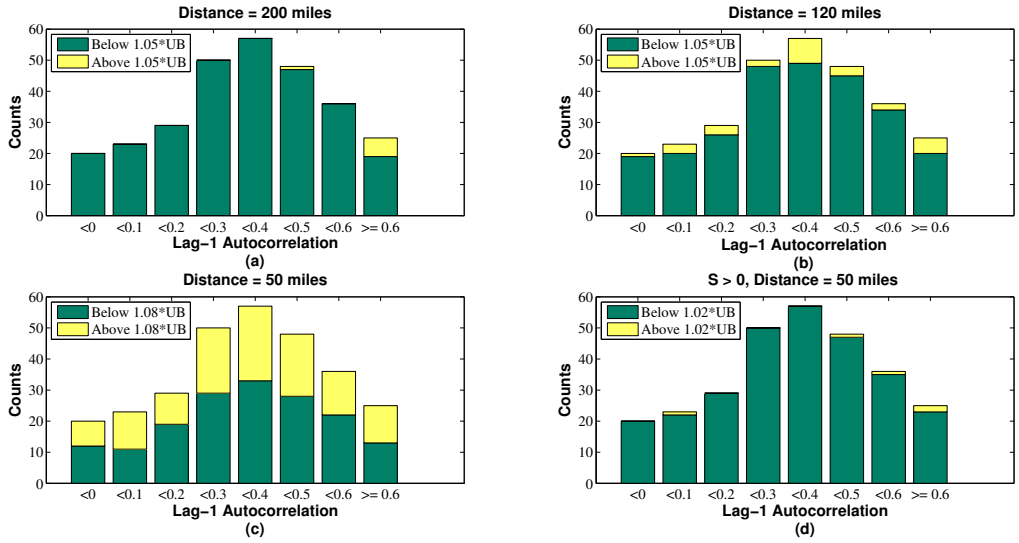

Figure EC.5 Distribution of lag-1 autocorrelation and the validity of the cost upper bounds.

• Last but not the least, Table EC.3 also shows that the combination of economic transmission line and ES can potentially significantly bring down the cost of building an ES-free line, as represented by the dash-dotted line in Figures EC.4 (a), (b) and (c). However, this cost-saving also depends on the fixed installation cost of ES, which we do not incorporate in these experiments. .

The above results and analysis suggest that the total variable cost and its associated transmission and ES capacity given by Section 5 are more effective estimates of the optimal values than those given by Section 4. This is because the former explicitly consider ES capacity cost and involve modeling inaccuracies that are canceling to each other. We thus also justify the heuristic approach to jointly planning ES-transmission network, as outlined in Subsection 5.3.

EC.4. Supporting Information for Section 6

EC.4.1. Experiment Settings of the Case Study in Subsection 6.1

To generate the problem instances for the case study in Subsection 6.1, we choose the sites of the potential wind farms identified in EPRI-DOE (2008), which also includes the data of wind outputs. The set of candidate junction sites J is the union of the set of the wind farms I and a set of additional locations in this region. We repeat the experiments by adding wind farms that are increasingly distant from Billings. In addition to the preceding parameters, we set the per-kW fixed transmission line cost $g = \$4.5 \times 10^5$/mile, fixed ES installment cost $h = \$1 \times 10^6$ and variable ES cost $r = \$0.5$/kWh. The experiments run on a Macbook Pro with 2.27 Ghz dual-core Intel Core i5 CPU and 8 GB memory. All instances are solved using CPLEX 12.3 called by YALMIP modeling language (Löfberg (2004)).

EC.4.2. Analysis of Technology Considerations in Subsection 6.2

The economic feasibility of each ES technology is significantly impacted by its cost-efficiency parameters, such as round-trip energy conversion efficiency, upfront construction cost and per-unit capacity cost. In Figure EC.4 in Section EC.3 we have already shown that systems cost as well as economic ES and transmission capacities are very sensitive to per-kWh cost. In what follows, we

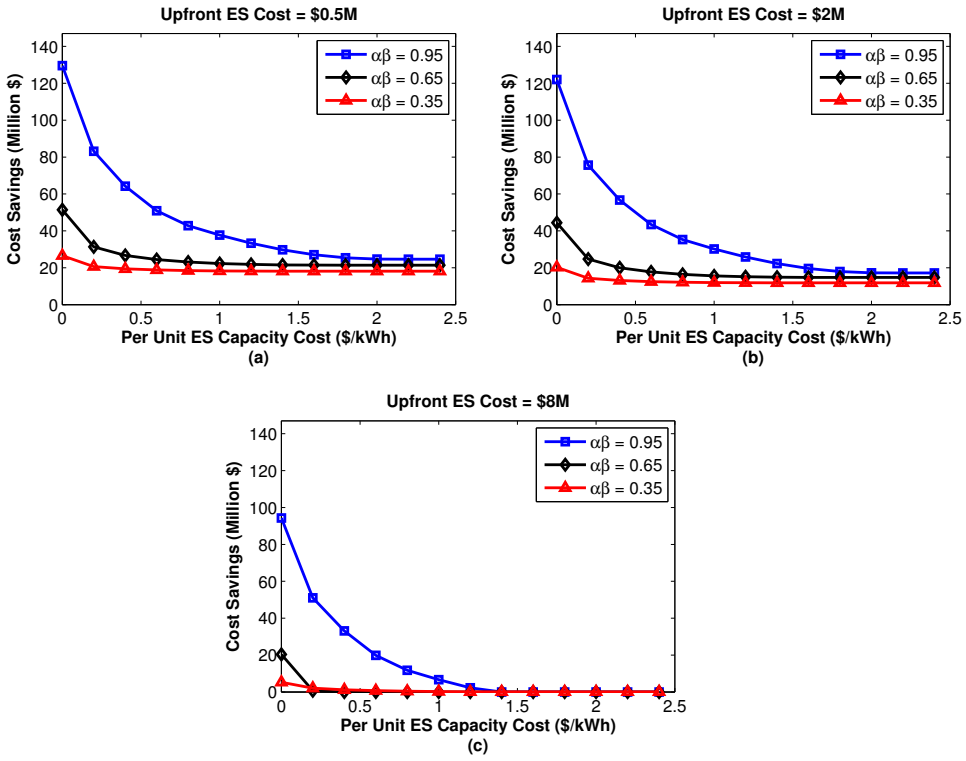

Figure EC.6 Impact of different ES technologies and their advancements on cost-savings.

evaluate how the cost-savings respond to the different values of these cost-efficiency parameters in our case study with 24 wind farms and 28 candidate junction sites. The cost-savings are benchmarked against a scenario without ES co-location. The results are shown in Figure EC.6, from which we draw several observations.

- Firstly, cost savings increase when (i) the round-trip conversion efficiency $\alpha\beta$ increases, (ii) per-unit ES capacity cost decreases and (iii) the fixed upfront ES cost decreases, as expected.
- Secondly, given upfront ES cost, when the per-unit ES capital cost is very high (e.g, $r \geq 1.6$ in our numerical test), cost-saving is small and cannot be effectively increased by enhancing conversion efficiency. This suggests that high per-unit capacity cost discourages adding ES capacity and it is not worth investing in improving the conversion efficiency, as in the case of lead-acid battery. However, when the per-unit ES capacity cost decreases (e.g, $r < 1.4$ in our numerical test), the cost-saving is more sensitive to reduction in per-unit ES capital cost when energy conversion is efficient (large $\alpha\beta$) than when it is inefficient (small $\alpha\beta$). As shown in Figure EC.6(b), for instance, when

the per-unit ES capital cost decreases from \$0.6/kWh to \$0.4/kWh, the cost saving increases by 30.1% with $\alpha\beta = 0.95$, compared with the cost saving increased by only 12.7% with $\alpha\beta = 0.65$. This suggests that more investment is desirable in improving conversion efficiency when the per-unit ES capacity cost has already been low, as in the case of compressed air energy storage systems.

- Thirdly, as for the upfront ES cost, it is interesting to see that we can achieve cost saving even with low conversion efficiency and high per-unit ES capacity cost (as shown in Figure EC.6(a) and (b)), until the upfront ES cost is too large to justify building any ES systems (as shown in Figure EC.6(c)). In the former case, the positive cost saving is mainly achieved by a small base ES capacity that salvages transmission capacity cost by smoothing short-term wind-out fluctuations.

The above observations collectively suggest that the cost-efficiency parameters jointly determine the effectiveness of using ES systems. Therefore, the planners need to carefully weigh these cost factors to properly size the ES systems, and the priority of ES R&D should be given to addressing the bottleneck cost factor, be it conversion efficiency or capital cost.