Preprint of the manuscript.
Full text available at: https://ieeexplore.ieee.org/abstract/document/8026163
Cite as: S. Majumder, S. A. Khaparde, A. P. Agalgaonkar, P. Ciufo, S. Perera and S. V. Kulkarni, "DFT-Based Sizing of Battery Storage Devices to Determine Day-Ahead Minimum Variability Injection Dispatch With Renewable Energy Resources," in IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 626-638, Jan. 2019.

DFT-Based Sizing of Battery Storage Devices to Determine Day-Ahead Minimum Variability Injection Dispatch With Renewable Energy Resources

Subir Majumder, *Graduate Student Member, IEEE*, Shrikrishna A. Khaparde, *Senior Member, IEEE*, Ashish Prakash Agalgaonkar, *Senior Member, IEEE*, Phil Ciufo, *Senior Member, IEEE*, Sarath Perera, *Senior Member, IEEE*, and S. V. Kulkarni, *Senior Member, IEEE*

***Abstract*—Renewable energy (RE) resources are non-dispatchable due to their intermittent nature, and battery storage devices (BSDs) play an important role to overcome their inherent variability. Therefore, for optimal operability, BSDs must be appropriately sized. Historical RE generation data can be used for sizing, with the objective to minimize the annualized planning cost. Application of high and low pass filters about a given cut-off frequency on the frequency spectrum of the historical generation data, calculated using discrete Fourier transform approach, segregate the fast and slowly varying components. The proposed methodology is based on 3σ principle and will ensure minimum injection of RE-generation variability into the grid for day-ahead scheduling with both fast and slowly varying components. The analysis shows that the batteries with the minimum unit capacity cost to throughput ratio provide minimum annualized planning cost for both slow and fast varying components. Determination of sizing of BSDs for a given cut-off frequency is numerically "costly" and to obtain the optimal cut-off frequency, a derivative-free mode-pursuing sampling method is applied. Exponential reduction in the daily injection of variability with increasing statistical significance in sizing is observed. Impact of unit-capacity cost to throughput-ratio on the sizing of BSDs is also studied.**

***Index Terms*—Battery storage devices, optimal sizing, discrete Fourier transform (DFT), mode-pursuing sampling (MPS) method.**

NOMENCLATURE

$\boldsymbol{A}$	A lower triangular matrix.
$\mathcal{C}$	Statistical capacity rating of batteries considering historical renewable energy (RE) generation data.
$\mathcal{C}'$	Modified statistical capacity rating of batteries to account for C-rating limit.
$\mathcal{C}_Z$	Statistical capacity rating of batteries considering the component Z of historical RE-generation data.
$\mathrm{C}^{\mathrm{K}}_{+}$, $\mathrm{C}^{\mathrm{K}}_{-}$	Effective capacity rating of storage '$+$' and '$-$' considering historical RE-generation of K^{th} day.
C_{rate}	C-rate of the battery.
d	Discount rate of an investment to calculate the annualized cost.
F_1, F_2	Low and high-frequency limits of fast Fourier transform.
$\mathcal{F}_S$	Throughput factor for battery storage device (BSD) type S.
Kb	Cost annualization factor of the BSDs.
k_h	Hour equivalent of day-ahead scheduling intervals.
Kp	Cost annualization factor of the power electronic (PE) converters.
$\mathcal{P}$	Statistical power rating of PE converters considering historical RE-generation data.
$\mathrm{G}^{\mathrm{K}}(t)$	Historical RE-generation at the time t of K^{th} day.
n	Index representing the days.
n_D	Capital recovery period.
P^{K}	Power rating of PE converters considering historical RE-generation of the K^{th} day.
$\mathrm{P}^{\mathrm{K}}_d(t)$	Combined dispatch from the RE-generators and BSDs at the time t of K^{th} day.
$\mathrm{P}^{\mathrm{K}}_b(t)$	Total power injected into the battery at the time t of K^{th} day.
$\mathrm{P}^{\mathrm{K}}_g(t)$	Total power injected into the BSD at the grid end of the BSD unit at the time t of K^{th} day.
PB	Combined dispatch from the RE-generators and BSDs.
PR	Historical Renewable Energy generation data.
PS	Power dispatched from the BSD.
$\mathrm{Q}^{\mathrm{K}}(t)$	Energy contained within a BSD at the time t of the K^{th} day.
$\mathcal{R}$	Statistical C-rating of BSD considering historical RE-generation data.

S. Majumder is with Department of Energy Science and Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India, and also with the Australian Power Quality and Reliability Centre, University of Wollongong, Wollongong, NSW 2522, Australia (e-mail: subirmajumder@iitb.ac.in).

S. A. Khaparde and S. V. Kulkarni are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India.

A. P. Agalgaonkar, P. Ciufo, and S. Perera are with the Australian Power Quality and Reliability Center, University of Wollongong, Wollongong, NSW 2522, Australia.

R^{K}	Average C-rating of the storage device for the K^{th} day.
η_{ch}, η_{dch}	Combined charging and discharging efficiency of the battery and the converter.
N_D	Total number of instants in a day for day-ahead scheduling.
N_Y	Total number of days in the historical data representing a year.
$\mathbf{Q}^{\mathrm{K}}$	A vector depicting energy contained within the battery.
S	Index represents different types of the batteries.
SOC'_{avg}	Statistically calculated desirable residual SOC of the battery.
t	Index representing the time instants in a day.
$\mathcal{T}$	Statistical throughput of the batteries considering historical RE-generation data.
T^{K}	Total throughput consumed by the BSD on K^{th} day.
U	Cost of PE-converters in \$/MW.
Y_S	Throughput available in the batteries of type S in MWh.
Z	Index representing low and high-frequency components of the historical RE-generation data.
$\beta_{S,Z}$	Binary variable representing the selection of storage types S for each component of Z.
Γ	Annualised investment cost.
Γ'	Modified annualized investment cost.
$\Delta\mathrm{SOC}$	Desirable depth of discharge of the battery.
Θ_S	Capacity cost of the batteries of type S in \$/MWh.
ϵ	A dimensionless small positive number.
$\Lambda(t)$	A sign variable symbolizing the charging and discharging of batteries.
Υ^{Z}	Annual throughput consumed by the BSD considering the component Z.

I. Introduction

RENEWABLE energy (RE) (such as, solar and wind energy) resources, though variable, exhibit non-zero correlation among themselves [1]–[3]. Low electricity demand at night is usually coupled with comparatively higher wind-energy production and zero solar power generation, and vice versa. However, the use of anticorrelation among RE generations and load are unreliable enough to meet continuous demand and production of electricity [2].

Energy storage devices help to reduce the inherent variability in various RE-generations. Use of storage devices to enable constant power injection into the grid to emulate base load generation is also well established in [4]–[8]. Since, RE-generation is stochastic in nature, combined constant RE-storage output can be obtained from the daily RE generation forecast. Hence, schedule of ESDs considering selling of "Only Renewable" [9], [10] energy will be a function of daily RE generation forecast. Therefore, RE generation and load are required to be assimilated, and the injection from ESDs would ensure the base-load generation objective.

Encapsulation of operational parameters in the planning process of storage devices is well established in [4]. Assuming that the emulated base-load generator is completely scheduled, transmission constraints are often ignored [5]–[7]. However, the generation and transmission expansion planning must be coordinated, and it can be shown that planning has a direct implication on the degree of operability [10]. Dehghan and Amjady [3] have suggested a coordinated robust transmission and storage expansion plan. In contrast, as a baseline, it is also important to study the sizing requirement assuming resilience of the transmission grid.

High power density, ramp rate, conversion efficiency, and capacity to weight ratio enable the use of battery based storage devices for electricity grid applications [11], without compromising the utility of other storage devices from electricity production point of view. Finite life of batteries, which is often ignored in the literature, imposes a limit on the use of batteries. Also, unit costs of batteries are still very high, and active research is underway for reduction of their cost [12] and improvement of their life. Batteries require power electronic (PE) converters interface to be able to be integrated into the AC electricity grid. But unlike the batteries, the converters are not needed to be frequently replaced. Therefore, an optimal investment would ensure overall minimization of the planning cost of the battery storage devices (BSDs).

Use of multiple types of storage devices for bulk energy storage in conjunction is well established in [13]–[16]. The existing literature suggests that the low-frequency component of RE generation must be associated with batteries with a finite number of cycles, while, super-capacitors or flywheels can be used in conjunction with high-frequency components. It is also important to note that although supercapacitors or flywheels are usually associated with large self-discharge rate [17], these are not often directly accounted while calculating the sizing of ESDs.

To facilitate the extraction of fast and slow varying component from the historical data, recently, signal-processing approaches [13], [14], [18]–[20] are actively pursued. Slow and fast varying components of the historical data are segregated out using either *discrete wavelet* based method [18] or *discrete Fourier transform* (DFT) approach [21]. A comparative study to present the application of DFT and discrete wavelet transform based methodology for sizing of the storage device is presented in [13]. As a part of the conventional signal processing approach, DFT technique converts the time domain signal into frequency domain – use band-pass filters to extract suitable frequency region obtained using DFT analysis – and convert these signal again into the time domain signal [19]. Capacity and power rating of the BSDs for each of slow and fast varying components are then calculated using Monte-Carlo approach [13], [14], [18], [20]. It is notable that because RE generation is stochastic, statistical calculation is essential to determine the sizing requirement. In the current context, in line with existing research findings, it can be expected that batteries with finite life-cycles will mitigate the slowly varying component in historical generation, and associated batteries with large life-cycle will mitigate the fast varying component.

Assuming similar ramp rate, conversion efficiency, physical footprint of BSDs, and zero self discharge rate, the following null hypothesis ($\mathcal{H}_0$) that constitutes of two parts can be stated: (i) batteries with comparatively large life-cycle will be used in conjunction with high-frequency component and vice versa, and (ii) operation of different type of batteries with various life-cycles in conjunction will be cost-effective. The total cost to be minimized will be the annualized value of the investment cost. Therefore, the goal is to calculate the cut-off frequency that minimizes the annualized cost. To achieve that sizing of batteries and converters for both slow and fast varying component defined relative to each cut-off frequencies (for low- and high-frequency component concerning a given cut-off frequency, see [15, Fig. 4]) are statistically calculated.

Presently, the choice of batteries for grid storage application is versatile. Although the cost of batteries is declining while possible physical life is improving, these improvements among various battery technologies may not be similar. Since the sizing of the BSDs is based on historical RE-generation data, degradation of BSDs per unit time is independent of the selection of BSD type, while, the replacement frequency will be exclusive to a particular battery type. Besides, active choice of battery for grid storage application does not remain same throughout the lifetime of a project. Therefore, it is intended to find out a strategy that ensures, once the sizing of the BSD is decided, the selection of battery type will be such that the annualized investment cost function remains independently at the minimum. Because the selection of battery type does not remain constant, investment cost will not remain constant, and therefore, the proposed methodology can be called as 'pay-as-you-go-plan'.

To summarize, in this paper, the following the research questions are asked:

i. Does batteries with a significant amount of storage cycle needs to be associated with fast storage cycle component of renewable generation, and vice versa?
ii. Does operation of multiple batteries is generic enough to reduce the annualized planning cost?
iii. Assuming resiliency of the grid as a baseline, what will be the implication of proposed "base-load emulation strategy" on the existing electricity network?

To answer these questions, following topics are included to describe the specific findings of this paper,

- A 'minimum variability injection' operation strategy;
- An optimal selection criterion of battery type that ensures minimum planning cost;
- A computational procedure to obtain the optimal cut-off frequency that minimizes the total planning cost of BSDs.

The rest of the paper is organized as follows. The 'minimum variability injection' optimization problem and statistical 3σ calculation of sizing of BSDs is discussed in Section II. Overall sizing methodology and the algorithm are discussed in Section III. The case study incorporating optimal sizing of BSDs and its impact on operational scheduling is discussed in Section IV. Section V concludes the paper.

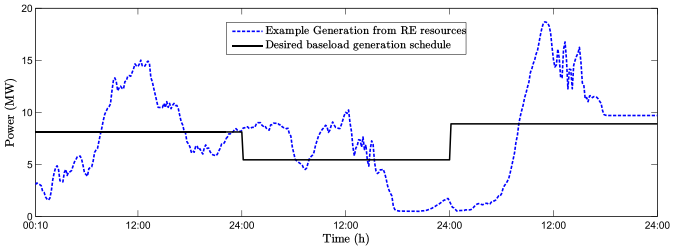


Fig. 1. Constant day-ahead generation schedule for a typical historical generation.

II. Sizing of BSDs for Minimum Variability Injection Generation Schedule

From the day-ahead operational point of view, it will be ideal if combined output power from RE-generation and BSDs are maintained at a constant level throughout the scheduled day. Other objectives can also be considered in this regard while keeping the independence of daily schedules. In order to achieve the independence, the average daily power output from a BSD must be zero [3], [10]. From the environmental point of view, it is essential to ensure *zero wastage*; and the equality condition in this regard is depicted in (1).

$$PR(t+(n-1)\cdot N_D)+PS(t+(n-1)\cdot N_D) = PB(n,t);\quad 1\le t\le N_D, 1\le n\le N_Y \tag{1}$$

Since the combined RE-BSDs output (PB) must remain constant, the equality constraint implies that the frequency response of PS should have been 180° phase apart with zero mean from the frequency response of PR. However, conversion efficiency (efficiency of PE-converter and batteries considered together) limits direct calculation of the power to be supplied from the BSDs. In addition, dispatchability and independence of each day-ahead schedules are the contradicting objectives. Therefore, if the sizing of BSDs are pre-calculated, constant dispatch from RE-generation and storage device while maintaining independent day-ahead schedules may not be ensured.

Fig. 1 shows the historical generation data from a combination of the wind and solar generator of 20 MW capacity to depict typical RE-generation data and its associated constant day-ahead dispatch schedule. The historical data obtained from [22] and [23] will be used as a part of case study.

A. Parameters for Sizing

If a battery is required to be integrated into an AC grid, a power electronic (PE) converter based interface is also essential; and the power rating of the PE-converter will be decided based on injectable (positive or negative) power into the batteries. While PE-converters need not be frequently replaced, with the finite physical and operational life of batteries, batteries are required to be frequently replaced throughout the project lifetime. The combined operational battery and PE-converters can be called as BSDs.

Batteries are required to be connected in series-parallel combination to maintain DC-bus voltage and satisfy the planned capacity constraint simultaneously. It can be assumed that the

total capacity requirement is an integer multiple of capacities of series connected batteries to maintain the DC-bus voltage, considering modularity of batteries. Expected operational life of batteries will be a function of current flowing through each series connected branch. Assuming charge balance circuitry is in place, the total current flowing from the battery will be equally divided among the parallel branches. According to available definition, the '1C-rate' of the battery is the required constant current output from batteries to discharge it within one hour altogether. Assuming DC-bus voltage is remaining constant, through simple manipulation, it can be shown that the C-rate of all batteries is constant, and the 1 C-rate is equal to the power to be extracted from the BSD to charge or discharge itself completely within 1 hour. Because the current derived from the BSD is not constant, the C-rate of the BSD (C_{rate}) with the capacity of $\mathcal{C}$ during the interval t can be defined as follows:

$$C_{rate} = \frac{|PS(t + (n-1) \cdot N_D)|}{\mathcal{C}}; \quad \forall t, n \tag{2}$$

Depth of discharge of each storage cycle (from Day-Ahead operational point of view) does not remain constant as well. Therefore, an alternate and simpler metric that signifies total number of cycles executed is applied, and it is termed as the *throughput* of batteries [24] is used in this problem to estimate capacity depreciation. Mathematically, throughput of a storage device (Υ) executed for n^{th} day, for the periods specified by t, with each period consists of k_h hours, can be defined by,

$$\Upsilon = k_h \sum_{\forall t} |PS(t + (n-1) \cdot N_D)| \tag{3}$$

Typically, each battery manufacturers, along with a capacity rating, specify the rated lifetime throughput (Y), which is proportional to the capacity rating of the storage device. The proportionality constant ($\mathcal{F}$) is unique for each type of batteries, and can be called as 'throughput factor.' For simplicity, 'throughput factor' can also be equal to the number of cycles that the battery can execute at its standard operating condition. Then rated executable throughput can be given by,

$$Y = \mathcal{F} \cdot \mathcal{C} \tag{4}$$

Because batteries have to absorb all the variability, C_{rate} and throughput expended by the battery in each operational day does not remain constant. However, to prevent fast degradation of batteries, C_{rate} must be limited, while, throughput expended to eliminate the variability determines replacement frequency of battery. It is notable that both C-rate and throughput spent depend on variability of historical data, and will play significant role in sizing.

References [25] and [26] indicates that the response time of most of the battery technologies is less than one second. Therefore, maximum physical ramp-rate can be executed by the BSDs will be limited by available charge within batteries, C_{rate} and ratings of PE-converters. Because ramp-rate limiting factors are already accounted in sizing, ramp-rate will not hinder the operation of the BSD.

B. Statistical Calculation for Sizing

Once the sizing of the BSD is determined, the objectives, (i) average output from RE-BSD combination to be constant throughout a day, and (ii) day-ahead scheduling to be independent of each other, as discussed, may not be satisfied simultaneously, because of asymmetric charging and discharging characteristics. Asymmetry arises because of the inherent inefficiency of battery and PE converters. The aim of reduction of injection of the variability in RE-generation into the grid can only be achieved by minimizing the squared sum of injection of variability, while total energy stored into a battery in a day is zero. The proposed objective can be called as the 'minimization of variability injection' criterion. Also, since day-ahead schedules are independent, average daily injection from BSD and RE-generator into the grid will not remain constant.

The sequence representing RE-generation on K^{th} day can be given by, $\mathrm{G^K}(t) = \{PR(t + (\mathrm{K}-1) \cdot N_D)\}$, $1 \le t \le N_D$. The optimization problem to obtain appropriate sizing of BSDs considering RE-generation data of the K^{th} day with an objective to minimize the injection of variability into the grid can be given by,

$$\min_{\mathrm{P}_d^{\mathrm{K}}(t), \mathrm{P}_b^{\mathrm{K}}(t), \mathrm{P}_g^{\mathrm{K}}(t)} \sum_{1 \le t \le N_D} \left(\mathrm{P}_d^{\mathrm{K}}(t) - \{\overline{\mathrm{P}_d^{\mathrm{K}}(t)}\} \right)^2 \tag{5}$$

subject to,

$$\mathrm{G^K}(t) - \mathrm{P}_g^{\mathrm{K}}(t) - \mathrm{P}_d^{\mathrm{K}}(t) = 0 \quad \forall t \tag{6}$$

$$\Lambda(t) = \frac{\mathrm{P}_g^{\mathrm{K}}(t)}{|\left(\mathrm{P}_g^{\mathrm{K}}(t)\right)| + \epsilon} \quad \forall t \tag{7}$$

$$\left(\mathrm{P}_b^{\mathrm{K}}(t) - \eta_{ch} \cdot \mathrm{P}_g^{\mathrm{K}}(t)\right) \cdot (1 + \Lambda(t)) \times \left(\mathrm{P}_b^{\mathrm{K}}(t) - \frac{\mathrm{P}_g^{\mathrm{K}}(t)}{\eta_{dch}}\right) \cdot (1 - \Lambda(t)) = 0 \quad \forall t \tag{8}$$

$$k_h \sum_{1 \le t \le N_D} \mathrm{P}_b^{\mathrm{K}}(t) = 0 \tag{9}$$

Ideally, $\Lambda(t)$ is a sign variable, where, $\Lambda(t) = +1$ symbolizes BSD is charging, and $\Lambda(t) = -1$ symbolizes that the BSD is discharging. ϵ is a small positive integer, that allows a smooth transition to monitor charging and discharging conditions of batteries.

The objective function (5) symbolizes the squared sum error of the variability, calculated for the daily average generation, to be minimized. Constraint (6) signifies the power balance equation and is required to be satisfied at each time intervals of a day-ahead schedule. Assuming combined charging and discharging efficiency of PE-converter and batteries to be η_{ch} and η_{dch}, constraint (8) calculates total power injected into the BSD. Constraint (9) signifies total charge stored within a BSD is zero ensuring independence of day-ahead schedules [3], [10]. If the residual energy contained within the battery be given by $\mathcal{C}_0$, the total energy/charge contained within the battery at time T can be given by, $\mathcal{C}_0 + k_h \sum_{t=1}^{T} \mathrm{P}_b^{\mathrm{K}}(t)$, where, $1 \le T \le N_D$.

Fig. 2 shows an example of different daily charging/discharging profile of batteries. Let SOC_{min} and SOC_{max}

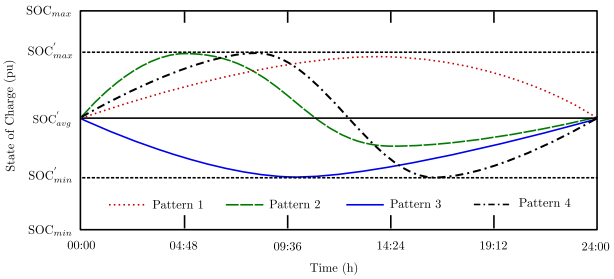


Fig. 2. Different SOC profiles to demonstrate the requirement of two-battery model. SOC$'_{max}$ and SOC$'_{min}$ represent maximum and minimum SOC executed by the battery for Pattern 4, calculated with respect to average SOC of SOC$'_{avg}$ for a given capacity.

be allowable minimum and maximum state of charge (SOC) of the battery, while SOC$_{avg}$ represent residual SOC to be retained at the beginning and end of a scheduling horizon. Also, let SOC$_{avg}$ is the algebraic mean of SOC$_{min}$ and SOC$_{max}$. Now consider, for example, for a given day, to ensure minimum variability injection, minimum and maximum state of charge of the battery can be given by SOC$'_{min}$ and SOC$'_{max}$ respectively (see Fig. 2); where SOC$'_{min}$ and SOC$'_{max}$ may not be symmetric about but are defined with respect to SOC$_{avg}$. On the other hand, SOC$'_{min}$ and SOC$'_{max}$ may vary for each of the scheduled day. To ensure predefined maximum depth of discharge (ΔSOC) of batteries, appropriate capacity rating of the battery can be calculated in such a way that $\Delta\text{SOC} = \text{SOC}'_{max} - \text{SOC}'_{min}$ for each scheduled day. To account for daily charging and discharging asymmetry the residual SOC may shift to SOC$'_{avg}$, where, SOC$'_{avg}$ may not be equal to the algebraic mean of SOC$'_{min}$ and SOC$'_{max}$. However, ensuring residual SOC to be at SOC$'_{avg}$, and DOD at ΔSOC will ensure complete utilization of the batteries for the complete historical data set, while maintaining desired depth of discharge.

A two battery model constituting of fictitious 'BSD +' and 'BSD −' is applied for sizing. Assuming $\mathcal{C}_0 = 0$, let 'BSD +' remains active when $k_h \sum_{t=1}^{T} \mathrm{P}_b^{\mathrm{K}}(t) > 0$, and 'BSD −' will be active when $k_h \sum_{t=1}^{T} \mathrm{P}_b^{\mathrm{K}}(t) < 0$. It is notable that total capacity can be obtained by simply adding capacities of 'BSD +' and 'BSD −'. Charge contained within the BSD ($\mathbf{Q}^{\mathrm{K}}$), capacity rating of 'BSD +' ($\mathrm{C}_+^{\mathrm{K}}$) and '−' ($\mathrm{C}_-^{\mathrm{K}}$), and throughput of the BSD (T^{K}) to eliminate the RE-generation variability considering the RE-generation of K^{th} day can be given by:

$$\mathbf{Q}^{\mathrm{K}} = \left\{ k_h \cdot \boldsymbol{A} \cdot \mathrm{P}_{\mathbf{b}}^{\mathrm{K}} \,\middle|\, \begin{array}{ll} A(i,j) = 0 & \text{if } i < j; \\ = 1 & \text{otherwise} \end{array} \right\} \tag{10}$$

$$\mathrm{C}_+^{\mathrm{K}} = \begin{cases} \frac{|\max\{\mathbf{Q}^{\mathrm{K}}\}|}{\Delta\text{SOC}} & \textit{if } \max\{\mathbf{Q}^{\mathrm{K}}\} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$\mathrm{C}_-^{\mathrm{K}} = \begin{cases} \frac{|\min\{\mathbf{Q}^{\mathrm{K}}\}|}{\Delta\text{SOC}} & \textit{if } \min\{\mathbf{Q}^{\mathrm{K}}\} < 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$\mathrm{T}^{\mathrm{K}} = k_h \sum_t \left| \mathrm{P}_b^{\mathrm{K}}(t) \right| \tag{13}$$

In (10), $\mathbf{Q}^{\mathrm{K}}$ represents a vector depicting energy contained within the battery with zero residual energy stored. It can be shown that $\max\{\mathbf{Q}^{\mathrm{K}}\} \geq 0$ and $\min\{\mathbf{Q}^{\mathrm{K}}\} \leq 0$. From modeling point of view, four scenarios are viable: (i) $\min\{\mathbf{Q}^{\mathrm{K}}\} < \max\{\mathbf{Q}^{\mathrm{K}}\}$; $\min\{\mathbf{Q}^{\mathrm{K}}\}, \max\{\mathbf{Q}^{\mathrm{K}}\} \neq 0$ (ii) $0 = \min\{\mathbf{Q}^{\mathrm{K}}\} < \max\{\mathbf{Q}^{\mathrm{K}}\}$, (iii) $\min\{\mathbf{Q}^{\mathrm{K}}\} < \max\{\mathbf{Q}^{\mathrm{K}}\} = 0$ and (iv) $\min\{\mathbf{Q}^{\mathrm{K}}\} = \max\{\mathbf{Q}^{\mathrm{K}}\} = 0$. Capacity calculation (11)-(12) accounts for the presented four scenarios, while for a given day, only one of these four components will be active. Equation (13) is based on calculation of throughput as shown in (3).

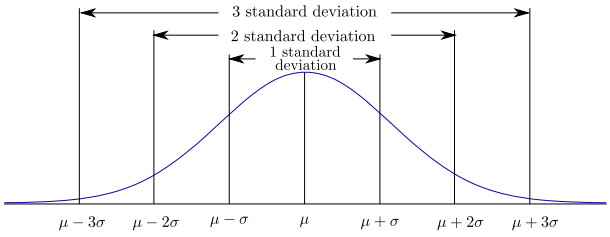


Fig. 3. Normal Distribution depicting 3σ interval.

Maximum of the power output from the battery determines power rating of the PE-converters, which can be given by:

$$\mathrm{P}^{\mathrm{K}} = \max\left\{ \left| \mathrm{P}_g^{\mathrm{K}}(t) \right| \right\} \tag{14}$$

The optimization problem formulated in (5)-(9) is required to be solved for the randomly selected days from historical time-domain data to calculate the sizing requirement. Sizing of the BSDs is statistically calculated with these daily storage requirements using 3σ principle described as follows.

Assuming, the given data set or observations follows normal distribution $\mathcal{N}(\mu, \sigma^2)$, the probability of occurrence of an event, $\mathcal{X}$, can be given by,

$$P\{|\mathcal{X} - \mu| < a\sigma\} = 2\Phi(\mathcal{X}) - 1, \qquad a > 0 \tag{15}$$

$\Phi(\mathcal{X})$ is the cumulative distribution function of normal distribution. Graphical representation of the normal distribution function along with its mean, μ and standard deviation σ are depicted in Fig. 3.

If $a = 3$, we get, $P\{\mu - 3\sigma < \mathcal{X} < \mu + 3\sigma\} = 0.99730$, which infers that the events $|\mathcal{X} - \mu| > 3\sigma$ are virtually impossible. Design of an experiment considering the events with $|\mathcal{X} - \mu| < 3\sigma$ embraces the event probability of 0.99730, which is called as 3σ principle [27].

In the current context, the schedule will be at risk, if the requirement exceeds design specifications. Selection of 3σ principle reduces the non-conformance probability to 0.0013. Reduction of the confidence interval will increase the non-conformance probability while reducing BSD ratings. In this work, the 3σ confidence interval is selected as a hard constraint. However, the effect of reduction of confidence interval on sizing of BSDs and associated performance of day-ahead schedule will also be studied.

The vectors consisting of power rating of the PE-converter ($\{\mathrm{P}^{\mathrm{K}} : \forall \mathrm{K}\}$), the capacity rating batteries ($\{\mathrm{C}_+^{\mathrm{K}} : \forall \mathrm{K}\}, \{\mathrm{C}_-^{\mathrm{K}} : \forall \mathrm{K}\}$), and daily throughput $\{\mathrm{T}^{\mathrm{K}} : \forall \mathrm{K}\}$ expense of the battery can be obtained for all the randomly selected days. Assuming the probability distributions of sizing requirement following normal distribution, with the estimated

mean and standard deviation of the sizing requirement for randomly selected days, the capacity rating batteries ($\mathcal{C}_+, \mathcal{C}_-$), power rating of PE converters ($\mathcal{P}$), and daily throughput expense of the battery ($\mathcal{T}$) can be calculated based on 3σ method, and can be given by:

$$\mathcal{C}_+ = \mu\big(\{\mathrm{C}_+^{\mathrm{K}} : \forall \mathrm{K}\}\big) + 3 \cdot \sigma\big(\{\mathrm{C}_+^{\mathrm{K}} : \forall \mathrm{K}\}\big) \tag{16}$$

$$\mathcal{C}_- = \mu\big(\{\mathrm{C}_-^{\mathrm{K}} : \forall \mathrm{K}\}\big) + 3 \cdot \sigma\big(\{\mathrm{C}_-^{\mathrm{K}} : \forall \mathrm{K}\}\big) \tag{17}$$

$$\mathcal{P} = \mu\big(\{\mathrm{P}^{\mathrm{K}} : \forall \mathrm{K}\}\big) + 3 \cdot \sigma\big(\{\mathrm{P}^{\mathrm{K}} : \forall \mathrm{K}\}\big) \tag{18}$$

$$\mathcal{T} = \mu\big(\{\mathrm{T}^{\mathrm{K}} : \forall \mathrm{K}\}\big) + 3 \cdot \sigma\big(\{\mathrm{T}^{\mathrm{K}} : \forall \mathrm{K}\}\big) \tag{19}$$

$\mu(\cdot)$ and $\sigma(\cdot)$ are functions depicting maximum likelihood estimate of mean and standard deviation (degree of freedom adjusted) respectively. Estimated $\mathcal{C}_+, \mathcal{C}_-$ are ≥ 0. Overall capacity rating of the battery ($\mathcal{C}$) and the residual SOC (SOC'_{avg}) can be given by:

$$\mathcal{C} = \mathcal{C}_+ + \mathcal{C}_- \tag{20}$$

$$\mathrm{SOC}'_{avg} = \frac{\mathcal{C}_-}{\mathcal{C}_+ + \mathcal{C}_- + \epsilon} \tag{21}$$

Because C-rate is dependent on the capacity rating of the BSDs, average C-rate of the BSDs for the K^{th} day is calculated and will be used for statistical calculation of overall C-rate of batteries. Calculation of the vector representing C-rate for randomly selected K^{th} day (R^{K}) and statistically calculated C-rate ($\mathcal{R}$) can be given as follows:

$$\mathbf{R}^{\mathrm{K}} = \overline{\left| \frac{\mathrm{P}_g^{\mathrm{K}}(t)}{\mathcal{C}} \right|} \tag{22}$$

$$\mathcal{R} = \mu\big(\{\mathrm{R}^{\mathrm{K}} : \forall \mathrm{K}\}\big) + 3 \cdot \sigma\big(\{\mathrm{R}^{\mathrm{K}} : \forall \mathrm{K}\}\big) \tag{23}$$

Capacity rating of batteries must be modified to $\mathcal{C}'$, without impacting SOC'_{avg}, so as to limit statistical C-rate ($\mathcal{R}$) to the upper bound of C-rate ($\mathcal{R}^{lim}$), to improve the life of batteries [28]. Calculation of $\mathcal{C}'$ is shown below:

$$\mathcal{C}' = \frac{\mathcal{C} \cdot \mathcal{R}}{\mathcal{R}^{lim}} \tag{24}$$

Statistically calculated capacity rating of the BSDs, will be used in the optimization problem described in the following section.

III. Calculation of Sizing of BSDs for Low and High Frequency Segments

The frequency response, $X(k)$ ($k \in \mathbb{Z}$) of a sequence representing the power to be injected into a BSD, $x(p)$ ($p \in \mathbb{Z}$) can be written as a sum of related complex exponential sequences [21]. This transformation is referred to as DFT, and can be depicted as:

$$X(k) = \sum_{p=0}^{N-1} x(p) W_N^{kp}; \qquad W_N = e^{-j(2\pi/N)} \tag{25}$$

'k' represents the frequency number in the discrete domain, and is related to time domain frequency by the Nyquist Frequency. N is the period of the historical data for the calculation of DFT. In the discrete domain, k is bounded by $[F_1, F_2]$ ($F_2 \geq F_1 \geq 0$).

Application of ideal low and high pass filter based on a particular cut-off frequency will separate out low and high-frequency component from the historical generation sequence. Band-limited time-domain signal is obtained from filtered frequency domain signals using inverse discrete Fourier transform (IDFT) method. The expression to get filtered time-domain signal can be given by,

$$\tilde{x}_{(\cdot)}(p) = \sum_{k=0}^{N-1} X(k) \cdot W_{(\cdot)} \cdot W_N^{-kp} \tag{26}$$

where, W_{LPF} and W_{HPF} are frequency response of low and high pass filter respectively. $\tilde{x}_{LPF}(p)$ and $\tilde{x}_{HPF}(p)$ represent slow and fast varying components of the historical data.

A. The Capital Recovery Factor of the Investment

Planning cost comprises of (i) cost of batteries, and (ii) cost of PE converters. Both of these expenses are required to be annualized using suitable discount rate (d) and recovery time (n_D) to obtain total annualized cost of BSDs. The annualization factor is called as capital recovery factor ($\mathrm{CRF}_b(d, n_D)$) and is defined as follows:

$$\mathrm{CRF}(d, n_D) = \frac{d(1+d)^{n_D}}{(1+d)^{n_D} - 1} \tag{27}$$

Multiplication of $\mathrm{CRF}(d, n_D)$ with the investment cost yields yearly recoverable cost of the project. The discount rate, 'd' is usually decided by investors. Life of PE converters often determines the duration of the project (assuming, PE-converters are 100% reliable within their lifetime). In contrary, physical life of batteries are smaller compared to PE-converters and driven by their throughput. Assuming the discount rate for recovering the cost of batteries to be very small, or, $(1+d)^{n_D} \approx (1 + n_D \cdot d)$, (27) reduces into:

$$\mathrm{CRF}_b(d, n_D) = d + \frac{1}{n_D} \approx \frac{1}{n_D} \tag{28}$$

The approximation is obtained considering $n_D \ll \frac{1}{d}$, which is true in general for batteries with physical life (n_D) of ≤ 5 years. If, the statistically calculated daily throughput to be executed by the batteries calculated using 3σ rule is given by Υ, and rated throughput executed is given by Y, with the throughput factor of $\mathcal{F}$, the recovery factor for the storage device (K_b) can be given by:

$$K_b = \mathrm{CRF}_b(d, n_D) = \frac{N_Y \Upsilon}{Y} = \frac{N_Y \Upsilon}{\mathcal{F}\mathcal{C}} \tag{29}$$

The recovery factor for the PE-converters (K_p) is constant, with both 'd' and 'n_D' being constant. It can be noted that both K_b and K_p are $\in \mathbb{R}^+$.

B. Optimum Choice of BSDs

The hypothesis ($\mathcal{H}_0$) was the use of different type of batteries with the various life-cycles in conjunction will be cost minimal, and batteries with large number of cycle will be utilized together with fast component of historical data set. In this regard, two types of batteries, $S \in \{A, B\}$, with the capacity cost of Θ_A and Θ_B \$/MWh and throughput factor of $\mathcal{F}_A$ and

$\mathcal{F}_B$ respectively are considered to be available at the beginning of planning horizon. Both $\mathcal{F}_A$ and $\mathcal{F}_B$ are intrinsic properties of BSDs, while, lifetime throughput ratings, Y_A and Y_B MWh respectively depends on the capacity rating of the BSDs as well.

It is notable that the fast Fourier transform (FFT) algorithm, used for calculation of DFT of the historical data set, is conjugate symmetric. The frequency response is discrete, and the frequency response is ranged within $[F_1, F_2]$. Therefore, the cut-off frequency (F) is required to vary within $[F_1, \frac{F_2+F_1+1}{2}]$ to capture the entire frequency spectrum; and, $F \in [F_1, \frac{F_2+F_1+1}{2}]$ ($F \in \mathbb{Z}^+$) divides the complex frequency response obtained from DFT into two components: (i) low frequency component $[F_1, F) \cup (F_2 + F_1 - F, F_2]$ and (ii) high frequency component $[F, F_2 + F_1 - F]$.

Let, $Z \in \{L, H\}$ indicate low and high-frequency components of the DFT spectrum about F. Application of IDFT on the low and high-frequency components of the DFT spectrum will generate the band-limited time-domain data. Capacity rating ($\mathcal{C}_Z$) and annual throughput (Υ_Z) of the BSDs, and PE-converter rating ($\mathcal{P}_Z$) for each of the components of Z are statistically calculated using 3σ rule. Either of the battery types can be selected for mitigating the variability of the low and high-frequency segments. The annualized cost function to be minimized can be written as:

$$\begin{aligned}\Gamma\big(F, \beta_{S,Z}\big) &= \sum_{Z\in\{L,H\}} \sum_{S\in\{A,B\}} K_b\, \mathcal{C}_Z \Theta_S \beta_{S,Z} \\ &\quad + K_p \cdot U\left(\sum_{Z\in\{L,H\}} \mathcal{P}_Z\right) \\ &= \sum_{Z\in\{L,H\}} \sum_{S\in\{A,B\}} \Theta_S \frac{\Upsilon_Z}{\mathcal{F}_S} \beta_{S,Z} \\ &\quad + K_p \cdot U\left(\sum_{Z\in\{L,H\}} \mathcal{P}_Z\right) \end{aligned} \tag{30}$$

For simplicity, let us assume that Θ_S and $\mathcal{F}_S$ are constant throughout the planning period. $\beta_{S,Z}$ is a binary variable ($\beta_{S,Z} \in \{0, 1\}$) indicating suitable selection of S, for each component of Z. If BSD type S is selected for segment Z, $\beta_{S,Z} = 1$; else, $\beta_{S,Z} = 0$. Since, at least one storage type is required to be selected for each component of Z; $\sum_{S\in\{A,B\}} \beta_{S,Z} = 1\ \forall Z$.

In the objective function (30), for a given F, Υ_Z and $\mathcal{P}_Z$ are constants; which will result into the cost of PE-converter to be constant. In the absence of cross-coupling term between low and high-frequency components of the cost of batteries in the objective function, the cost of batteries of the low and high-frequency segment must be independently at their minimum. For a given F, the objective function becomes,

$$\min \sum_{S\in\{A,B\}} \Upsilon_Z \frac{\Theta_S}{\mathcal{F}_S} \beta_{S,Z} \qquad \forall Z \tag{31}$$

Because Υ_Z are constants and ≥ 0, minimum investment cost for a given F depends only on minimum $\frac{\Theta_{(\cdot)}}{\mathcal{F}_{(\cdot)}}$ ratio.

Theorem 1: For minimization of total investment cost with multiple available BSDs, a storage device with minimum 'unit cost to throughput factor' ratio will only participate into the optimal mix independent of frequency segment considered.

Theorem 1 transforms the objective function (30) into:

$$\Gamma'(F) = \frac{\Theta_{S^*}}{\mathcal{F}_{S^*}} \left(\sum_{Z\in\{L,H\}} \Upsilon_Z\right) + K_p \cdot U\left(\sum_{Z\in\{L,H\}} \mathcal{P}_Z\right) \tag{32}$$

S^* represents optimum storage type based on minimum $\frac{\Theta_{(\cdot)}}{\mathcal{F}_{(\cdot)}}$ ratio.

Because of limited throughput of the batteries, in contrast to the converters, the batteries are required to be frequently replaced. In addition, $\frac{\Theta_{(\cdot)}}{\mathcal{F}_{(\cdot)}}$ may not remain constant until the end of life of batteries. But, as the PE-converter is designed at the beginning of the planning period for both segments, the operating frequency F has also become constant. In the present case, the planning cost remains at its minimum level, if batteries with minimum $\frac{\Theta_{(\cdot)}}{\mathcal{F}_{(\cdot)}}$ ratio is selected for successive replacement. And so, this planning methodology can be termed as 'pay-as-you-go-plan'.

The design specification of the BSDs used in the objective function depicted in (32) is calculated based on the statistical calculation presented in Section II. Therefore, the cut-off frequency that minimizes the objective function will be determined by the historical data set.

C. Solution Methodology

The cost-minimization objective function is dependent on the statistically calculated sizing of BSD. For each of the cut-off frequencies, a large number of secondary 'minimization of variability injection' optimization problem was required to be solved to determine the statistical sizing. Therefore, calculation of sizing of batteries and PE-Converters at each cut-off frequency is computationally intensive, or 'costly,' and therefore can be treated as a 'black box.'

The solution space of the optimization problem is discrete in nature, and derivative of the cost function is not readily available. Therefore, mode-pursuing sampling (MPS) method [29] has been used to solve the optimization problem. MPS method is efficient and robust for the optimization problem involving computationally expensive black-box functions with discrete search space. MPS is a derivative-free optimization algorithm and uses random-discretization based sampling method [30] to calculate new samples. This approach statistically evaluates a large number of samples near the global optima, ensuring faster convergence.

Algorithm 1 and 2 depict the solution methodology using MPS method. If sizing is to be calculated considering different historical RE-generation data, they are required to be assimilated (PM) and fed into the algorithm as an input. Other parameters to be fed into the algorithm are, search space specified by frequency range F_1 and F_2, randomly sampled days (n_R) for statistical calculation of the ratings of the BSD, n_P different frequencies in each iteration, n_C successive iteration with no change in optima and the convergence tolerance, Δ.

Essentially, Algorithm 1 generates a set of frequencies (with n_P number of elements) from the discrete search space obtained from frequency range F_1 and F_2, and calls

Algorithm 1: Algorithm to Calculate Optimal BSD Rating

Data: F_1, F_2, n_P, n_C, n_R, Δ
Result: ${\mathcal{C}_Z}^*$, ${\Upsilon_Z}^*$, ${\mathcal{P}_Z}^*$
• Randomly select n_R number of days, D from a uniform distribution;
• Randomly select n_P different cut-off frequencies ($F \in [F_1, \frac{F_2+F_1+1}{2}]$) without repetition, and calculate the cost function, Γ'_F using Algorithm 2 (extreme points are forcefully included to validate the proposed hypothesis), considering randomly selected D;
• Create a set n consisting of calculated costly data points, and their corresponding functional values;
• Set temporary variable $flag = 0$, to track convergence of the MPS algorithm;
• Set an arbitrary high cost, Γ_{opt};
while $flag \leq n_C$ **do**
  − Probabilistically select n_P different cut-off frequencies (F) without repetition, excluding existing points using MPS algorithm;
  − Based on the set of costly new data points, calculate the cost function, Γ'_F using Algorithm 2, considering randomly selected D;
  − Append the set of costly data points n with new samples and associated cost functions;
  − Find the minimum of the calculated samples, Γ^N_{opt};
  if $\left|\frac{\Gamma_{opt}-\Gamma^N_{opt}}{\Gamma_{opt}}\right| \leq \Delta$ **then** $flag = flag + 1$;
  else $flag = 0$;
  if $\Gamma_{opt} \leq \Gamma^N_{opt}$ **then**
    $\Gamma_{opt} = \Gamma^N_{opt}$;
  end
end
• Calculate Γ^*; ${\mathcal{C}_Z}^*$; ${\Upsilon_Z}^*$; ${\mathcal{P}_Z}^*$ corresponding to the cut-off frequency with the cost function value Γ^N_{opt};

Algorithm 2: Algorithm to Statistically Calculate Cost and Rating of BSDs for a Given Cut-Off Frequency, F

Data: PM, F, D
Result: $\mathcal{C}_Z$, Υ_Z, $\mathcal{P}_Z$, Γ'_F
• Apply low and high pass filters with respect to the cut-off frequency F on the DFT sequence;
• Apply IDFT on these signals to obtain band-limited time-domain signals;
for $\forall D$ **do**
  − Calculate capacity rating of the BSD, power rating of the converters, and throughput of the BSDs;
end
• Statistically calculate $\mathcal{C}^Z$, Υ^Z, $\mathcal{P}^Z$;
• Using expression (22), and (23) calculate $\mathcal{R}^Z$;
if $\mathcal{R}^Z > \mathcal{R}^{lim}$ **then**
  − Recalculate storage sizing using expression (28) such that it conforms to average C-rate limit;
end
• Calculate the cost function $\Gamma'(F)$ using equation (32);

TABLE I
SPECIFICATION OF THE BSD FOR PLANNING

Parameters	Specifications
Battery Specifications	\$120/kWh (cost), 1000 (throughput) [26]
PE-Converter Specification	\$800/kW (cost), 20 years (life) [14]
Charging and discharging efficiency	$\eta_{ch}, \eta_{dch} = 0.80$ [14]
ϵ	1×10^{-10}
C-rate threshold, $\mathcal{R}_{lim}$	3
Depth of discharge, ΔSOC	0.8
Discount rate of PE Converter, d	3%
N_D, N_Y	144, 365

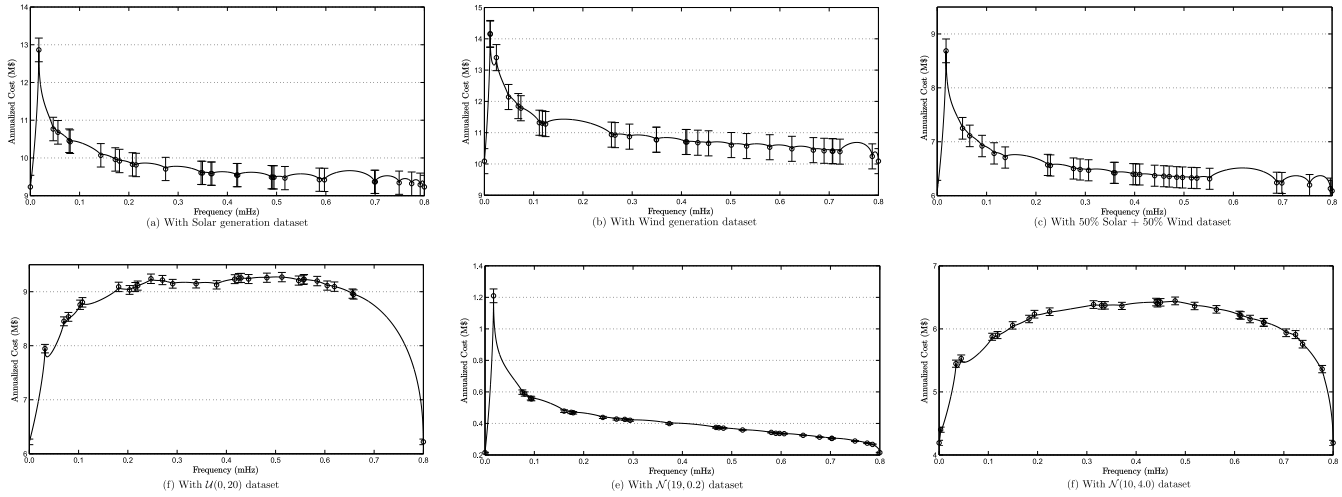
Algorithm 2 to statistically calculate the sizing of BSD for each of the generated frequency. Algorithm 2 uses assimilated historical data PM to convert it into low and high frequency segment with respect to the cut-off frequency obtained from Algorithm 1 using (25)-(26). Sizing requirements are then calculated using (5)-(14) for randomly selected days of each segment of historical data. Statistical sizing specifications are then calculated using (16)-(24) and fed it back to Algorithm 1. Algorithm 1 checks the convergence, then, either generate new data points or terminates indicating that the problem is converged. The problem is assumed to be converged, if the optimal solution does not change for n_C number of successive interval. If the problem is not converged, in each iteration, n_P additional frequencies are probabilistically selected as a part of MPS algorithm.

n_R random numbers, each representing a day, were sampled from an uniform distribution. Statistical sizing calculations are carried out, for each of these sampled days. Because, sizing of low- and high-frequency segments were required to be independently calculated, a total of $2n_R$ secondary optimization problem was solved at each cut-off frequency to evaluate the sizing of batteries and PE-converters to mitigate the uncertainty in the historical data. It is also notable that similar set of days are required to be used at each cut-off frequency for statistical sizing calculation.

IV. CASE STUDY

Yearly historical generation data set of similar resolution can be assimilated, and the frequency spectrum was obtained using DFT. The Nyquist rate of the frequency signal was 0.8 mHz (resolution of the dataset is 10 minutes). Installed capacity of both the wind-generating and the solar-generating site was 20 MW. As proven earlier, BSDs with minimum $\frac{\Theta_{(\cdot)}}{\mathcal{F}_{(\cdot)}}$ ratio was selected at the beginning of planning; and the associated specification is given in Table I.

In addition to the available historical wind [22] and solar [23] generation data sets for the wind and solar generation site of 20 MW capacity each, an additional hypothetical data set was created using the algebraic summation of 50% Solar and 50% Wind generation. Additionally, a uniformly distributed data set ($\mathcal{U}(0, 20)$) was generated with maximum and minimum of 0 and 20 respectively. Two normally distributed data sets, with mean and standard deviation of 19 and 0.2 ($\mathcal{N}(19, 0.2)$), and 10 and 4.0 ($\mathcal{N}(10, 4.0)$) respectively are also generated, while truncating them to limit within 0 and 20. All of these artificially created data sets along with original Wind and Solar generation data represents various scenarios



(a) With Solar generation dataset (b) With Wind generation dataset (c) With 50% Solar + 50% Wind dataset

(f) With $\mathcal{U}(0, 20)$ dataset (e) With $\mathcal{N}(19, 0.2)$ dataset (f) With $\mathcal{N}(10, 4.0)$ dataset

Fig. 4. Variation in annualized cost at different cut-off frequency for minimum variability injection.

TABLE II
PARAMETERS OF THE MPS METHOD

Randomly sampled days, n_R	120
Costly points selected per iteration, n_P	5
Iteration limit for no change in optima, n_C	4
Convergence tolerance, Δ	10^{-3}

of RE-generation and has been used to identify the applicability of the hypothesis while testing the proposed methodology. Length of each of the data set is selected in such a way that they can represent historical RE-generation data with the resolution of 10 minutes, while the data set is available for 365 days representing a year.

A. Sizing of BSDs to Mitigate Variation in RE-Generation

Fig. 4 depicts the variation in the annualized cost function for the discrete cut-off frequencies for the generated scenarios. The parameters used for MPS algorithm is given in Table II. Also, the discrete frequencies are converted into time domain using the Nyquist rate. Since, sizing of the BSD for each scenario were calculated only at finite number cut-off frequencies as a part of MPS method, Fig. 4 was obtained by fitting a cubic spline. In each of the scenarios, the minimum value of cost function was found out to be lying at both the extremes of the cost function, indicating use of single BSD embracing complete frequency spectrum will be cost optimal.

Because, given a cut-off frequency, sizing of BSD is calculated based on randomly selected samples, the solution obtained will be uncertain. The variation in the cost with finite sample size is assumed to follow the normal distribution, and standard error in the total cost has also been presented in Fig. 4 by error bar. Estimation errors are calculated using the methodology presented in [31]. It is known that the estimation error asymptotically reduces to zero with the selection of a large number of days as samples. It is also notable in Fig. 4 that in each of the scenario at each cut-off frequency, coefficient of variation (CV) of the total cost is very small.

Two observations are notable in the current context: (i) the battery with significant life-cycle, not necessarily to be used to mitigate high-frequency component of RE-generation variation, it is rather unit cost to throughput factor ratio which will dictate the sizing, and (ii) application of two BSDs in conjunction (maybe two batteries of similar type) such that either of them can be used to mitigate slow- and fast- varying components need not result in most economic planning. The cost curves signify that the sizing of BSDs considering the complete frequency spectrum will be cost minimal. However, because of the use of a finite number of scenarios, in this work, "evidence of absence" [32] can not be theorized. Therefore, use of multiple batteries (can be applied to storage devices in general) in conjunction will require special treatment, and sizing can be calculated on case by case basis.

Table III shows the sizing solutions of batteries and PE-converters for cost-minimal planning in each of the scenarios. Overall sizing requirement considering the complete frequency spectrum in each of the scenarios are shown. It is also notable that in each of the cases residual SOC level need not remain at 0.5 to ensure the DOD limit.

It is evident that the sizing and indirectly the planning cost will be proportional to the variability within the historical generation data. The second moment has been used to represent the variability within the stochastic data set. Fig. 5 simultaneously represent cost per unit capacity of each of the data set signified by their second moment with varying statistical significance selected for sizing, and linear fit obtained using least squares estimation (LSE) between cost per unit capacity with the second moment of each of the data set with statistical significance as a parameter. Both Table III and Fig. 5 indicate that sizing requirement increases with increasing variability within the dataset. Table III also shows that CV because of finite sampling is very small.

From Fig. 5, it can be concluded that a linear increase in variability will linearly increase the unit planning cost. The coefficient of the linear term is the function of other parameters, such as statistical significance selected for sizing, unit cost to throughput ratio of batteries, unit cost of converters and

TABLE III
COMPARISON OF THE RATING OF THE BSDS

	$\mathcal{C}$ (MWh)				Υ (MWh)				$\mathcal{P}$ (MW)				SOC
	μ		σ		μ		σ		μ		σ		
	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error	
With Solar generation only	63.00	1.33	14.59	0.95	100.90	2.98	32.67	2.11	12.07	0.22	2.43	0.16	0.50
With Wind generation only	53.85	3.79	41.47	2.69	91.43	3.87	42.43	2.75	9.39	0.29	3.20	0.21	0.41
With 50% Solar and 50% Wind generation	41.75	2.29	25.12	1.63	69.24	1.89	20.66	1.34	7.30	0.17	1.81	0.12	0.46
With $\mathcal{U}(0, 20)$ data set	17.04	0.57	6.29	0.41	117.81	0.50	5.52	0.36	10.93	0.05	0.53	0.03	0.52
With $\mathcal{N}(19, 0.2)$ data set	0.57	0.02	0.24	0.02	3.77	0.02	0.21	0.01	0.58	0.01	0.08	0.01	0.52
With $\mathcal{N}(10, 4.0)$ data set	11.82	0.45	4.91	0.32	75.82	0.38	4.18	0.27	10.28	0.05	0.59	0.04	0.50

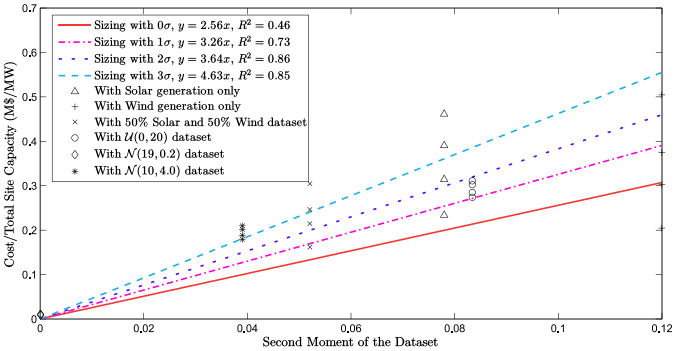


Fig. 5. A linear fit of the cost per unit capacity of BSDs with different moment of the data sets.

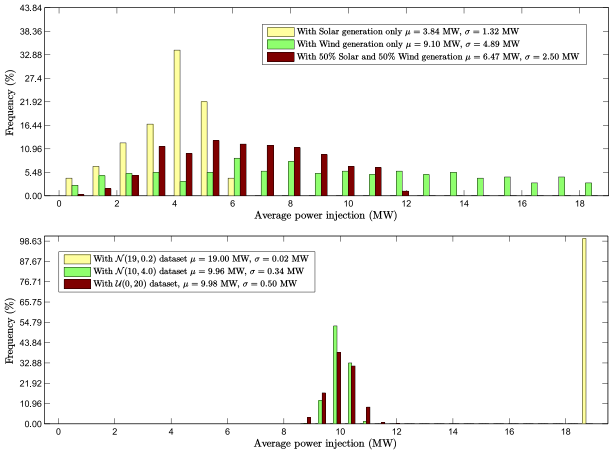


Fig. 6. The histogram of average daily net injection into the grid with sizing according to 3σ principle.

various recovery factors. It is also notable that if fitted using LSE, the coefficient of the linear term also linearly increases with increasing variability. In overall, characteristic curve can be given by, $y = 0.68x(p + 3.77)$, with $R^2 = 0.79$; where, x is the variability within the dataset measured using the second moment, and p is the statistical significance used for sizing measured using the standard deviation. Moreover, $p = -3.77$ indicates $y = 0$; symbolizing if sizing is calculated based on approximately -3σ, BSDs will not be required and injection of variability into the grid will not be bounded.

B. Benefits of Sizing Using 3σ Principle

To analyze the utility of sizing using 3σ principle, given the capacity rating, throughput rating, the residual SOC of batteries and power rating of the converters, minimization

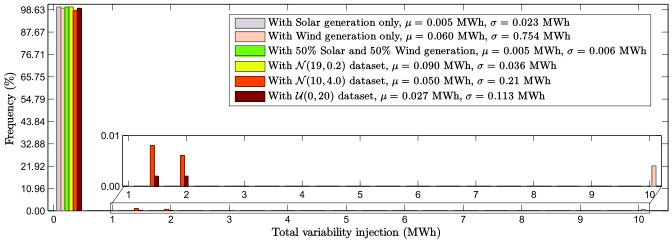


Fig. 7. The histogram of deviation in average daily net injection into the grid.

of injection of RE-generation variability into the grid shown in (5)-(9) was carried out. Histogram of average daily injection and total daily variability injected into the grid are presented in Fig. 6 and Fig. 7. The benefits of using 3σ principle can be observed in Fig. 7, which indicates, little to no injection of variability into the grid. However, independence of daily scheduling does not ensure the invariability of average daily injection into the grid. Fig. 6 represents that depending on the selected data set, the daily injection has a large standard deviation. Because the second moment for $\mathcal{N}(19, 0.2)$ data set is ≈ 0, sizing requirement is also ≈ 0.

From Fig. 6 it is notable that both average and standard deviation of daily power injected into the grid is very low for solar generator only operation, while it is very high for wind generators. With equal mixing solar and wind generation, the mean of average daily injection has been reduced by 28.90% compared to the wind only operation, while the standard deviation of average daily injection has been reduced by 48.88%. This way, with an optimal mix of various RE generation, the mean capacity utilization can be increased while decreasing the variability in average daily injection, paving the way towards an RE only grid.

C. Benefits Regarding Reduction of Variability With Increasing Statistical Significance in Sizing

Statistical 3σ principle is selected as a hard constraint in this work. However, it is expected that with reduction in statistical significance in terms of standard deviation for sizing the variability injection will also be increased.

Fig. 8 shows that the total daily variability injection is exponentially decreasing with increasing statistical significance. While it can be shown that increasing statistical significance would linearly increase the total planning cost with no change in the selection of cut-off frequency. Therefore, a proportional increase in the planning cost will incur an exponential reduction in the daily injection of variability into the grid, and so,

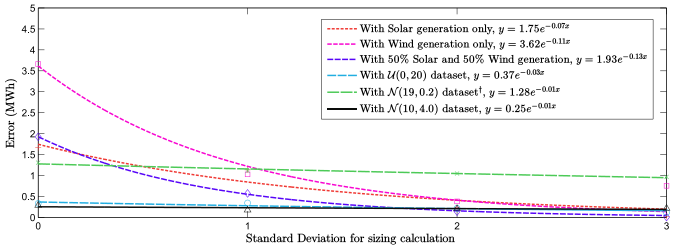


Fig. 8. Injection of variability with increasing statistical significance.

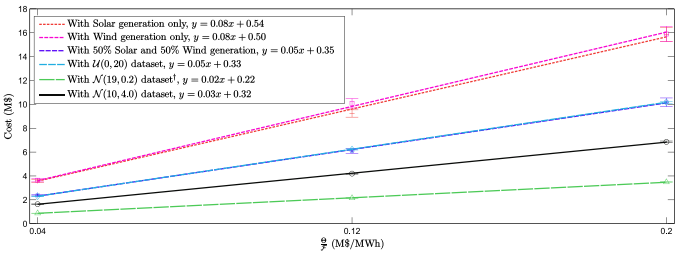


Fig. 9. Impact of *unit cost to throughput factor* on annualized planning cost.

the returns from increasing statistical significance are diminishing. Nevertheless, if statistical significance is not selected as a hard constraint, the statistical significance itself is required to be a criterion to study the economic feasibility, and will be encountered in the future studies.

D. Impact of Unit Cost to Throughput Factor on the Planning Cost

Because unit cost to throughput ratio itself is variable throughout the planning horizon, and also, the batteries with least unit cost to throughput factor ratio will be required to be selected for successive replacement, it is important to study the impact of the ratio on the cut-off frequency and the total cost.

Fig. 9 shows that the total cost linearly increases with increasing unit cost to throughput factor ratio. The objective function, can be represented as, $\Gamma'(F) = x(\sum_{Z\in\{L,H\}} \Upsilon_Z) + K_p \cdot U(\sum_{Z\in\{L,H\}} \mathcal{P}_Z)$, and the co-efficient can be directly obtained from total annual throughput $\sum_{Z\in\{L,H\}} \Upsilon_Z$, and annualized cost of converters $K_p \cdot U(\sum_{Z\in\{L,H\}} \mathcal{P}_Z)$. x is the unit cost to throughput ratio, as defined earlier.

Fig. 9 is obtained by independently calculating the annual throughput and the annualized cost, and R^2 are ≥ 0.9. It can be shown that *if a sequence is the linear sum of its subsequences, and the global optimal point does not shift with a change in non-negative weights, then each of its components resides independently at their global optima, and vice-versa.* Therefore, because both total annual throughput and annualized cost of converters, remains constant independent of the $\frac{\Theta}{\mathcal{F}}$ ratio, and cut-off frequency of the objective functions lies at the extreme points of the search space, the optimal value of the components also independently lies at both the extremes and can be verified by further simulations.

Therefore, for the given the data set in all the scenarios, and 3σ statistical significance, it is always optimal to consider the complete frequency spectrum in sizing. Because the change in the ratio has no impact on the cut-off frequency, it can be expected that lifetime total planning cost remains at the minimum with "pay-as-you-go" plan. In this case, because the annual throughput and annualized cost are independent of the ratio, increase in CV will be observed with increasing $\frac{\Theta}{\mathcal{F}}$ ratio.

†Observations are scaled up by a factor of 10, to improve the clarity of the figures.

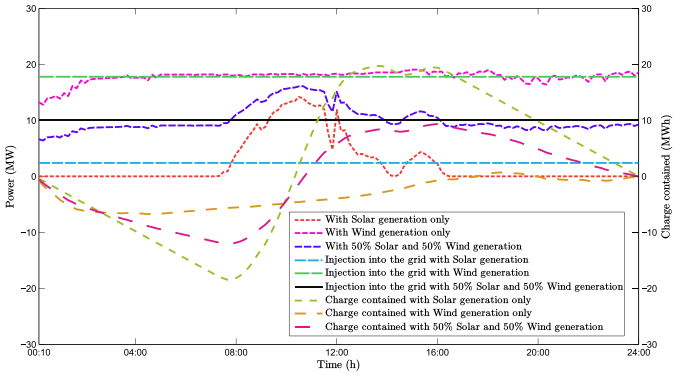


Fig. 10. Injection into the grid, the BSDs, and charging condition of batteries.

E. Impact of Wind and Solar Generation in the Day-Ahead Operation of the Grid

Fig. 10 shows an example of the hourly operation of RE-BSD combination for an arbitrarily selected day. The calculation of the residual SOC level has been presented earlier and will differ based on the historical data, while invariability of initial and final SOC ensures independent day-ahead operation. As expected from Fig. 7, interval wise injection into the grid in both the scenarios remains constant for the scheduled day, while, the order of statistical significance in selecting the appropriate sizing of BSDs would have a direct impact in the injection of variability into the grid. Asymmetry of charging and discharging profile around the residual SOC is also notable.

F. Real-Time Operation

Proposed 'minimum variability injection' methodology can be suitably incorporated in real-time balancing. It is notable that both day-ahead and real-time operations can complementarily eliminate the variability. In real-time operation, the primary focus is to alleviate the variability introduced by forecast error and not to mitigate the variability as a part of 'minimum variability injection' scheme. Assuming forecast error to follow a given distribution, mathematically, it can be assumed that the prediction error to follow stochastic random walk (Brownian movement). The sizing solution for real-time operation can be estimated from this assumption and can be a part of future work. It is notable that the intended formulation still assumes transmission network to be resilient enough to absorb all the power injected into the node.

V. Conclusion

An optimal sizing methodology of BSDs to mitigate the variability of RE generation was discussed in this paper. For

each of the cut-off frequencies, obtained from the DFT, the sizing of batteries and PE-converters were statistically calculated based on the historical RE-generation data. Since constant daily schedule and their independence can not be simultaneously satisfied, the design objective selected was the minimization of squared sum of variability, while, independence of daily schedule was a hard constraint. A 'cost to throughput factor' criterion was found for the initial selection of batteries that ensures minimum planning cost.

Because of the discrete solution space and computation intensive statistical method for sizing, to alleviate rigorous calculation involved in the exhaustive search to find out a global optima, MPS based method was applied. Various artificially generated scenarios were used to calculate the sizing for comparison purpose. In all the scenarios MPS show that the cut-off frequency for the minimum cost is located at either of the extreme, while, the use of multiple battery types in conjunction will be costlier. The proposed method is generic, but, the different historical data set will give a different result. It has also been found that the average daily injection into the grid is not constant, while, the histogram of total daily variability injection signifies that the proposed sizing methodology based on 3σ principle can successfully be used for base-load generation. For the data set used, the average injection with solar generation is less variable compared to that with the wind generation. Therefore, it can be concluded that an optimal mix of the wind and solar generation is desirable. It is also shown that the residual SOC level depends on the historical data set while ensuring the DOD limit. Exponential reduction in the injection of variability can be observed if statistical significance is varied while determining the sizing. The planning cost linearly increases with increase in 'cost to throughput factor' ratio. It was found that for the given datasets, and 3σ statistical significance, it is always optimal to consider complete frequency spectrum in sizing, and the proposed 'pay-as-you-go' plan attains the global minima.

REFERENCES

[1] L. Baringo and A. J. Conejo, "Correlated wind-power production and electric load scenarios for investment decisions," *Appl. Energy*, vol. 101, pp. 475–482, Jan. 2013.

[2] P. E. Bett and H. E. Thornton, "The climatological relationships between wind and solar energy supply in Britain," *Renew. Energy*, vol. 87, pp. 96–110, Mar. 2016.

[3] S. Dehghan and N. Amjady, "Robust transmission and energy storage expansion planning in wind farm-integrated power systems considering transmission switching," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 765–774, Apr. 2016.

[4] M. Zidar, P. S. Georgilakis, N. D. Hatziargyriou, T. Capuder, and D. Škrlec, "Review of energy storage allocation in power distribution networks: Applications, methods and future research," *IET Gener. Transm. Distrib.*, vol. 10, no. 3, pp. 645–652, Feb. 2016.

[5] X. Y. Wang, D. M. Vilathgamuwa, and S. S. Choi, "Determination of battery storage capacity in energy buffer for wind farm," *IEEE Trans. Energy Convers.*, vol. 23, no. 3, pp. 868–878, Sep. 2008.

[6] D. L. Yao, S. S. Choi, K. J. Tseng, and T. T. Lie, "A statistical approach to the design of a dispatchable wind power-battery energy storage system," *IEEE Trans. Energy Convers.*, vol. 24, no. 4, pp. 916–925, Dec. 2009.

[7] Q. Li, S. S. Choi, Y. Yuan, and D. L. Yao, "On the determination of battery energy storage capacity and short-term power dispatch of a wind farm," *IEEE Trans. Sustain. Energy*, vol. 2, no. 2, pp. 148–158, Apr. 2011.

[8] M. Korpaas, A. T. Holen, and R. Hildrum, "Operation and sizing of energy storage for wind power plants in a market system," *Int. J. Elect. Power Energy Syst.*, vol. 25, no. 8, pp. 599–606, Oct. 2003.

[9] E. D. Castronuovo and J. A. P. Lopes, "Optimal operation and hydro storage sizing of a wind–hydro power plant," *Int. J. Elect. Power Energy Syst.*, vol. 26, no. 10, pp. 771–778, 2004.

[10] S. Majumder and S. A. Khaparde, "Revenue and ancillary benefit maximisation of multiple non-collocated wind power producers considering uncertainties," *IET Gener. Transm. Distrib.*, vol. 10, no. 3, pp. 789–797, Feb. 2016.

[11] X. Luo, J. Wang, M. Dooner, and J. Clarke, "Overview of current development in electrical energy storage technologies and the application potential in power system operation," *Appl. Energy*, vol. 137, pp. 511–536, Jan. 2015.

[12] D. L. Anderson, "An evaluation of current and future costs for lithium-ion batteries for use in electrified vehicle powertrains," M.S. thesis, Nicholas School Environ., Duke Univ., Durham, NC, USA, 2009.

[13] H. Bitaraf, S. Rahman, and M. Pipattanasomporn, "Sizing energy storage to mitigate wind power forecast error impacts by signal processing techniques," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1457–1465, Oct. 2015.

[14] J. Xiao, L. Bai, F. Li, H. Liang, and C. Wang, "Sizing of energy storage and diesel generators in an isolated microgrid using discrete Fourier transform (DFT)," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 907–916, Jul. 2014.

[15] Y. Liu *et al.*, "Sizing a hybrid energy storage system for maintaining power balance of an isolated system with high penetration of wind generation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3267–3275, Jul. 2016.

[16] T. Senjyu, D. Hayashi, A. Yona, N. Urasaki, and T. Funabashi, "Optimal configuration of power generating systems in isolated island with renewable energy," *Renew. Energy*, vol. 32, no. 11, pp. 1917–1933, 2007.

[17] A. Buchroithner *et al.*, "Decentralized low-cost flywheel energy storage for photovoltaic systems," in *Proc. Int. Conf. Sustain. Energy Eng. Appl. (ICSEEA)*, Jakarta, Indonesia, 2016, pp. 41–49.

[18] Q. Jiang and H. Hong, "Wavelet-based capacity configuration and coordinated control of hybrid energy storage system for smoothing out wind power fluctuations," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1363–1372, May 2013.

[19] Y. V. Makarov, P. Du, M. C. W. Kintner-Meyer, C. Jin, and H. F. Illian, "Sizing energy storage to accommodate high penetration of variable energy resources," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 34–40, Jan. 2012.

[20] I. N. Moghaddam and B. Chowdhury, "Optimal sizing of hybrid energy storage systems to mitigate wind power fluctuations," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Boston, MA, USA, 2016, pp. 1–5.

[21] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[22] *National Renewable Energy Laboratory, Offshore_Sites_00002-06996*. Accessed: Jun. 2014. [Online]. Available: http://www.nrel.gov/electricity/transmission/eastern_wind_dataset.html

[23] *National Renewable Energy Laboratory, Solar Power Data for Integration Studies, al-pv-2006.* Accessed: Dec. 2016. [Online]. Available: http://www.nrel.gov/grid/solar-power-data.html

[24] M. Y. Nguyen, D. H. Nguyen, and Y. T. Yoon, "A new battery energy storage charging/discharging scheme for wind power producers in real-time markets," *Energies*, vol. 5, no. 12, pp. 5439–5452, 2012.

[25] Electrical Energy Storage Project Team, "White paper: Electrical energy storage," Int. Electrotech. Commission, Geneva, Switzerland, 2011. [Online]. Available http://www.iec.ch/whitepaper/pdf/iecWP-energystorage-LR-en.pdf

[26] G. Albright, E. Jake, and S. Al-Hallaj, "A comparison of lead acid to lithium-ion in stationary storage applications," Chicago, IL, USA, AllCell Technol. LLC, White Paper, 2012.

[27] V. Czitrom and P. D. Spagon, *Statistical Case Studies for Industrial Process Improvement*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 1997, pp. 341–342.

[28] Y. H. Jung, C. H. Lim, and D. K. Kim, "Graphene-supported $Na_3V_2(PO_4)_3$ as a high rate cathode material for sodium-ion batteries," *J. Mater. Chem. A*, vol. 1, no. 37, pp. 11350–11354, 2013.

[29] L. Wang, S. Shan, and G. G. Wang, "Mode-pursuing sampling method for global optimization on expensive black-box functions," *Eng. Optim.*, vol. 36, no. 4, pp. 419–438, 2004.

[30] J. C. Fu and L. Wang, "A random-discretization based Monte Carlo sampling method and its applications," *Methodol. Comput. Appl. Probab.*, vol. 4, no. 1, pp. 5–25, 2002.

[31] S. Ahn and J. A. Fessler, "Standard errors of mean, variance, and standard deviation estimators," Elect. Eng. Comput. Sci. Dept., Univ. Michigan, Ann Arbor, MI, USA, Tech. Rep. TR 413, Jul. 2003.
[32] C. Sagan, "The Demon–Haunted world: Science as a candle in the dark," 1997.

Subir Majumder (GS'16) received the B.Tech. degree in electrical engineering from the West Bengal University of Technology, India, in 2012, and the M.Tech. degree in energy systems engineering under the joint collaboration between the Indian Institute of Technology Bombay, India, and the University of Wollongong, Australia, where he is currently pursuing the Ph.D. degree.

His research interest includes power system optimization.

Shrikrishna A. Khaparde (SM'91) received the Ph.D. degree from the Indian Institute of Technology Kharagpur, India, in 1981.

He is a Professor with the Department of Electrical Engineering, Indian Institute of Technology Bombay, India. He has co-authored a book entitled *Computational Methods for Large Sparse Power Systems Analysis: An Object Oriented Approach* (Norwell, MA, USA: Kluwer, 2001). He is a member of Advisory committees to Maharashtra electricity regulatory commission, India and Indian Energy Exchange, India. His current research areas are smart grids, distributed generation, renewable energy policies, restructured power systems, and CIM Implementation in India.

Ashish Prakash Agalgaonkar (SM'13) received the B.E. degree in electrical engineering and the M.E. degree in electrical power system from the Walchand College of Engineering, Sangli, India, in 1997 and 2002, respectively, and the Ph.D. degree in energy systems engineering from the Indian Institute of Technology Bombay, India, in 2006.

He is currently a Senior Lecturer with the University of Wollongong. His research interests include planning and operational aspects of renewable and distributed generation, power system reliability, microgrids, electricity markets, and system stability.

Phil Ciufo (SM'07) received the B.E. (Hons.), M.E. (Hons.), and Ph.D. degrees in electrical engineering from the University of Wollongong, Australia, in 1990, 1993, and 2002, respectively.

He is currently an Associate Professor with the University of Wollongong. His research interests include modeling and analysis of power distribution systems and ac machines, advanced distribution system automation, and power quality.

Sarath Perera (SM'12) received the B.Sc. (Eng.) degree in electrical power engineering from the University of Moratuwa, Sri Lanka, in 1974, the M.Eng.Sc. degree from the University of New South Wales, Australia, in 1979, and the Ph.D. degree in electrical engineering from the University of Wollongong, Wollongong, Australia, in 1988.

He is currently a Professor and the Technical Director of the Australian Power Quality and Reliability Center, University of Wollongong.

S. V. Kulkarni (SM'08) received the B.E. degree from Bombay University, Bombay, India, in 1988, and the M.Tech. and Ph.D. degrees from the Indian Institute of Technology Bombay, India, in 1990 and 2000, respectively.

He is currently a Professor with the Electrical Engineering Department, Indian Institute of Technology Bombay. He has co-authored the book entitled *Transformer Engineering: Design, Technology, and Diagnostics* (CRC). His research interests include analysis and diagnostics of power transformers, computational electromagnetics, and distributed generation. He was a recipient of the Young Engineer Award from the Indian National Academy of Engineering and the Career Award for Young Teachers from All India Council for Technical Education. He is a fellow of the Indian National Academy of Engineering.