Received: 8 April 2021 | Accepted: 9 June 2021
DOI: 10.1002/2050-7038.13010

RESEARCH ARTICLE

WILEY

Machine intelligent forecasting based penalty cost minimization in hybrid wind-battery farms

Harsh S. Dhiman[1] | **Dipankar Deb**[2] | **S. M. Muyeen**[3] | **Ajith Abraham**[4]

[1]Department of Electrical Engineering, Adani Institute of Infrastructure Engineering, Ahmedabad, Gujarat, India
[2]Department of Electrical Engineering, Institute of Infrastructure Technology Research and Management, Ahmedabad, Gujarat, India
[3]School of Electrical Engineering Computing and Mathematical Sciences, Curtin University, Perth, Western Australia, Australia
[4]Machine Intelligence Research Labs (MIR Labs), Scientific Network for Innovation and Research Excellence, Auburn, Washington, USA

Correspondence
Harsh S. Dhiman, Department of Electrical Engineering, Adani Institute of Infrastructure Engineering, Ahmedabad, Gujarat 382421, India.
Email: harsh.dhiman@aii.ac.in

Summary
Modern-day hybrid wind farm operation is fundamentally dependent on the accuracy of short-term wind power forecasts. However, the inevitable error in wind power forecasting limits the power transfer capability to the utility grid, which calls for battery energy storage systems to furnish the deficit power. This manuscript addresses a wind forecasting based penalty cost minimization solution for hybrid wind-battery farms. We choose six wind farm sites (three offshore and the other three onshore) to study machine intelligent forecasting based solutions and compare the performance of a wavelet-Twin support vector regression (TSVR) based wind power forecasting model with ε-Twin support vector regression, Random forest, and Gradient boosted machines, for penalty cost minimization. We access the penalties that arise as power imbalances along with the battery system's cost. We find that TSVR based wind power forecasting method results in a minimum global operational cost for all the wind farm sites under study.

KEYWORDS
battery energy storage systems (BESS), machine learning (ML), penalty cost, wind farms, wind forecasting

1 | INTRODUCTION

Various studies concerning optimal allocation of power units of for planning and operation have been carried out.[1] Multi-objective optimization-based studies have demonstrated that it is convenient to design a hybrid power system comprising wind farm as a base power plant. However, the inherent intermittency in wind resource makes it difficult for grid operators to plan their dispatch in a congested power system network.[2] Stochastic nature of wind speed makes it difficult for the wind farm operators to plan their dispatch. This stochastic nature can be typically modeled in the form of a nonstationary and nonlinear time-series. Statistical models used to predict wind speed often fail as they lack the ability to trace the nonlinear trend in a wind speed time-series. An efficient wind farm operation is based on achieving economies of scale depending on the nature of the associated electricity market.

With increased wind power penetration globally, the importance of the time-scale in wind power forecasting has increased. Statistical model-based forecasting methods allow wind speed prediction over a large horizon, while the variations in wind speed on a smaller time-scale like that of 1-minute to 10-minutes give a deeper insight into various dynamic phenomena occurring in atmospheric boundary layer.[3] Wind turbine and farm control in modern power plants also rely on the preview wind speed information to be available in the range of 1-second to 10-minutes.

List of Abbreviations: ANN, Artificial Neural Network; BESS, Battery Energy Storage System; GBM, Gradient Boosted Machines; QPP, Quadratic Programming Problem; RF, Random Forest; SVR, Support Vector Regression; TSVR, Twin Support Vector Regression.

Commonly used induction control and yaw control techniques are used when preview information is available at a higher frequency, thus allowing a precision equipment like Light Detection and Ranging (LIDAR) to effectively control the wind turbines.[4] A choice of forecasting model either based on the statistical model or LIDAR based scanning is essential for wind farm control.

In terms of energy trading, minute-scale forecasting allows wind farm operators to obtain short-term forecasts where the imbalance is created due to the stochastic nature of wind speed. Minute-scale forecasts are also necessary for providing ancillary services, secondary or tertiary reserve for balancing capacity of a large pool of utilities. Grid stability is dependent on the balance between generation and demand. With fluctuations in wind power, it is important for wind generation companies to participate in the reserve power market according to their generations. It has also been observed that the spot price in an electricity market increases if the wind power forecasts are made for a longer horizon.[5,6] Previous studies have shown the usage of thermal power generation in order to control imbalances by wind power generation.[7] In the present case, a hybrid wind farm is utilized in order to reduce the costs incurred by a transmission system operator (TSO) for dispatch windows with imbalances. A TSO is responsible for maintaining coordination between generation and demand and also to handle any contingencies that may jeopardize the system security. Incurring a penalty cost to wind generation companies will encourage the usage of accurate wind forecasting methods.

Time horizon based forecasting is another important aspect in wind power forecasts.[8,9,10] A Majority of related research is now progressing towards machine learning algorithms that have the ability to derive a regression function to be used on new data. Wind speed data for a particular regime can be assessed in different time intervals like very short-term, short-term, and medium-term forecasting. Individual machine learning methods like Artificial neural networks (ANN), Support vector regression (SVR), Extreme learning machine (ELM), and Gaussian process regression are used in tandem with signal preprocessing methods like wavelet transform and empirical mode decomposition. Signal decomposition techniques segment the time-series into sub-series which are forecasted and then an aggregation provides the series. These techniques form the basis of a hybrid forecasting model where the advantages of individual models are synthesized to arrive. Among these methods, Yuan et al provides a hybrid forecast using auto-regressive fractionally integrated moving average (ARIFMA) and least-square support vector regression for short-term wind power prediction.[11]

Machine-learning algorithms have significant edge over classical statistical models like ARMA, ARIMA, and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH). Furthermore, signal decomposition based techniques have shown significant improvement in prediction accuracy in tandem with optimization of hyper-parameters.[12,13,14,15,16] In this work, we study Support vector regression (SVR) and its variants along with random forests and gradient boosted machines as regression models for short-term wind speed forecasting. SVR is derived from the core idea of support vector machines where a regressor function is determined based on a nonlinear mapping from input space to a higher dimension space. Similar to SVR, Least square support vector regression (LS-SVR), Twin support vector regression (TSVR), and ε-Twin support vector regression (ε-TSVR) are used for regression analysis.[17,18,19,20,21] A hybrid model based on wavelet transform and variants of SVR is discussed in Dhiman et al.[22] All these methods discussed use large historical data to train their respective algorithms. Computational complexity also plays an important role in formulating the hybrid models for wind speed forecasting. LS-SVR based model uses equality constraints in formulating the optimization problem and takes less time than conventional SVR. In terms of prediction accuracy, tuning of hyper-parameters results in a lower error which can be beneficial from reserve capacity point of view when dealing with imbalances in a hybrid wind farm. For random forests and gradient boosted machines, parameter tuning in form of the number of trees, and learning rate results in an improved prediction.[23]

For hybrid wind farms, decision-making strategies like Simple additive weighting (SAW), Technique for the order of preference by similarity to ideal solution (TOPSIS), and Complex proportional assessment (COPRAS) are used to find the preferred alternative.[24] Conceptually, a hybrid wind-battery farm focuses on penalty cost minimization achieved with accurate wind power forecasting. A wind farm operator ensures that the power schedules are available ahead in time in order to facilitate optimal power flow in a power system network.[25] Wind power is variable in nature and hence the excess or deficit in wind power accounted via forecasting can be treated with reserve power capacity. Battery energy storage systems are commonly employed for this purpose. However, due to its high investment cost and limited lifetime, modeling its operation with grid-connected power system is challenging. Hence, we formulate an analysis that takes into consideration the degree of accuracy with which wind power forecasts can be made. The formulation of penalty cost(s) incurred is purely as per the actual and predicted wind powers. The details of these costs are provided in the next section. The associated decision-making idea for a hybrid wind farm consists of tangible and intangible parts that determine the most preferred strategy for a wind farm operator at a given time. However, this is not the case when it comes to a particular dispatch window where different approaches incur different costs, depending on the type of forecasting method used. This manuscript aims to find the best forecasting approach among various possible ones.

The decision and control aspect in a hybrid wind farm faces continuous challenges, and with increasing wind power penetration the need for an efficient solution that ultimately aims to increase the annual energy production, is to be looked at closely.[26] Simultaneously, the advancement in the field of machine learning, accurate wind forecasts will potentially help the wind generation companies to operate in a higher margin of profit and dispatch maximal power to the utility grid.[27] Thus, in order to arrive at a holistic solution aimed at achieving an economic benefit to the wind farm operator, an integrated solution based on machine intelligent forecasting is proposed in this manuscript which stems from the efficacy of machine learning models to perform well for wind speed datasets as discussed in Dhiman and Deb.[21] The main contributions of this work can be summarized as follows:

1. A penalty cost minimization for wind-battery farm using wavelet-Twin support vector regression (TSVR) and a collective objective aimed to increase battery lifetime.
2. Evaluate various penalty costs incurred to farm operator due to wind power imbalances. Costs in the form of penalties are evaluated for the proposed forecasting methods for wind farm sites. This analysis gives direction to new avenues in the field of hybrid wind-battery farms.
3. The aim is not only to achieve maximum profit by implementing a superior forecasting method but also to enhance the lifetime of battery storage systems. Hence, the battery is taken into consideration by assessing the consecutive battery state changes from charging to discharging.
4. A common cost metric for the forecasting technique under consideration is evaluated under two different scenarios: (a) static which acts as a baseline case, and (b) dynamic, a more realistic case where the farm operator may not always have the choice to operate at an optimal battery threshold point is evaluated.

The remainder of the manuscript is structured as follows: Section 2 discusses the modeling of an integrated solution based on short-term power forecasts followed by Section 3 where Twin support vector regression is discussed along with its mathematical formulation. Section 4 presents results and discussions based on real-time data for wind farm sites followed by a concluding section. The basis of selection of these four machine learning methods is based on the ability of these regressors to perform equally well for wind speed forecasting problem as discussed in Dhiman et al.[22] The extension of these four machine learning techniques for penalty cost minimization also ascertains the ability of these techniques to be leveraged in short-term wind power dispatch.

2 | MODELING AN INTEGRATED SOLUTION FOR WIND FARMS

A grid-connected hybrid wind-battery system deals with specific challenging issues during its lifetime. Installation of Wind turbines in a given land area occurs as per the micro-siting results for optimal power capture.[28,29,30] However, in real-time, the market-driven scenarios do not guarantee an adequate power transfer to the grid. In the next subsections, we discuss the decision-making and battery health improvement aspects in hybrid wind farms. The operator tries to minimize the penalties incurred due to power generation imbalances and battery health assessment as per the number of consecutive charging-discharging instances.

2.1 | Decision-making environment

The decision-making environment for a hybrid wind farm involves an appropriate operational strategy that results in a minimum cost over a period of time. A wind farm operating with its aim to deliver power to the grid often has auxiliary power support in the form of BESS which operates either in charging or discharging mode. For deploying the integrated solution, we consider a multi wind farm dispatch scenario where every wind farm under consideration tries to minimize its cost incurred. The Operational cost incurred to a wind farm operator is purely based on the available forecast schedules. The availability of wind power forecast for a definite time horizon is an important factor that must match with the market timing where the energy balances are cleared between 5 minutes to 6 hours. For wind generation stations participating in short-term electricity markets, accurate forecasting method results in lower penalty costs and also auxiliary costs for the TSO. Figure 1 illustrates an infographic regarding different types of penalty costs considered for modeling an integrated solution for a hybrid wind-battery farm.

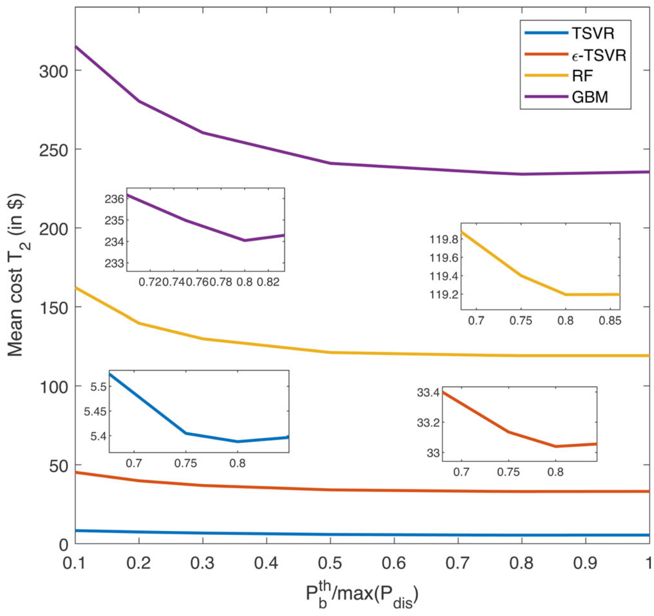

FIGURE 1 Framework under decision-making environment of integrated solution

Consider a wind farm operator for farm A who faces four penalty costs in the form of T_1, T_2, T_3, and T_4 as described in Figure 1. The wind farm operator working closely with the electricity market has to provide the wind power generation forecasts ahead in time in order to schedule a dispatch requiring auxiliary support to satisfy the grid requirements. Thus, for modeling an integrated solution, we take up the tangible part of the decision-making and analyze the same under different forecasting schemes.

Under this environment, a wind farm operator tries to minimize the cost incurred due to an imbalance in the forecasted wind power generation and the actual one. Imbalances are dealt with in terms of penalty which the farm operator minimizes by implementing a superior forecasting method. We determine the four penalty costs for wind farm **A** in each dispatch window. The wind farm operator of farm **A** follows the strategy resulting in minimum cost. The total cost is an aggregated value in each dispatch window, given as

$$J^* = \sum_{i=1}^{k} \min\{\mathbf{T_{1,i}}, \mathbf{T_{2,i}}, \mathbf{T_{3,i}}, \mathbf{T_{4,i}}\}, \tag{1}$$

where $T_{j,i}$ refers to the T_jth cost for the ith dispatch window, $i = 1, 2, ...k$ and $j = 1, 2, ..., 4$. The above cost J^* differs as the forecasting scheme changes. Individually, these costs can be expressed as follows

- Cost T_1 is based on the fact that hybrid wind-battery farm operator pays penalty when actual wind power is lower than the forecasted one. Consider $\hat{p}_i$ as the predicted wind power and p_i be the actual one, for k_a such scenarios, penalty cost is

$$T_1 = \beta_w \sum_{i=1}^{k_a} (\hat{p}_{A,i} - p_{A,i}), \tag{2}$$

where β_w is cost incurred to operator for \$ per 1 kW of shortfall in predicted wind power.

- Cost T_2 considers a scenario where the power from battery energy storage is used up to a certain threshold. Given there are m_s such instances where the deficit power exceeds threshold, this cost can be expressed as

$$T_2 = \begin{cases} \sum_{i=1}^{m_s} \left(\zeta_s \left(p_b^{\text{th}} - \hat{p}_{A,i} + p_{A,i}\right) + \delta_x \left(\hat{p}_{A,i} - p_i\right)\right), \\ \text{if } p_{A,i} - \hat{p}_{A,i} > p_b^{\text{th}}, \\ \beta_w \sum_{i=1}^{m_s} \left(p_{A,i} - \hat{p}_{A,i}\right), \text{ Otherwise,} \end{cases} \tag{3}$$

where ζ_s denotes per kW penalty in $ when deficit exceeds battery threshold.

- Cost T_3: Consider multiple wind farms, for example three in this case. The farms in proximity to each other can deliver deficit wind power in lieu of the wind farm under consideration. Hence, a penalty corresponding to the borrowed power must be paid. Let $p_{B,i}$, $\hat{p}_{B,i}$, $p_{C,i}$ and $\hat{p}_{C,i}$ be the measured and predicted powers for wind farms B and C respectively, the operator of farm A pays a penalty that is expressed as

$$T_3 = \begin{cases} \alpha_z \sum_{i=1}^{l_z} \left(p_{B,i} - \hat{p}_{B,i}\right), \text{ if } p_{B,i} > \hat{p}_{B,i}, \\ \alpha_z \sum_{i=1}^{l_z} \left(p_{C,i} - \hat{p}_{C,i}\right), \text{ if } p_{C,i} > \hat{p}_{C,i}, \\ \delta_x \sum_{i=1}^{l_z} \left(\hat{p}_{A,i} - p_{A,i}\right), \end{cases} \tag{4}$$

where α_z and δ_x penalty cost for per kW of borrowed power and penalty for power delivered through battery energy storage.

- Cost T_4 considers a scenario where the penalty cost is paid for power delivered entirely by battery energy storage. This cost can be expressed as

$$T_4 = \delta_x \sum_{i=1}^{u_l} \left(\hat{p}_{A,i} - p_{A,i}\right). \tag{5}$$

Evaluation of this cost is also important from a transmission system operator's point of view as any imbalance caused in wind power generation reflects on the ancillary cost to be borne by the TSO. Under the current scenario, we evaluate costs using Twin support vector regression, ε-Twin support vector regression, Random forest, and Gradient boosted machines. Such supervised machine learning based regression models use historical data. This generalization capability of machine learning models avoid over-fitting and thus increases the accuracy of the models which ultimately reduce the penalty on a wind farm operator. Cost T_2 which uses power from battery to be dispatched during imbalance, is only paid if the deficit exceeds a battery threshold P_b^{th}. The variation of this cost with different battery threshold naturally affects the overall cost incurred. A farm operator chooses the battery threshold which leads to minimum cost. With respect to the penalty costs, the objective is to minimize them (penalty costs) by incorporating an accurate machine intelligent forecasting technique (such as TSVR, epsilon-TSVR, RF, and GBM). The accuracy of the forecasts affects the magnitude of these penalty costs. Thus, in this work, the forecast that yields minimum cost common cost (CC), is found optimal for the operation of hybrid wind-battery farms. In terms of the profit, the penalty costs can be minimized in a particular dispatch window by carrying out an accurate wind power forecast.

2.2 | Battery state change cost under wind power forecasting

Recent wind farms are equipped with BESS to provide auxiliary support during imbalances in power generation. Often a battery system goes through a period of successive charging and discharging which can be harmful for its lifetime. Battery health is of prime importance as the need for charging and discharging BESS may arise due to constraints placed on the transmission network. Imbalance markets may also need to be in a state to accept excess wind power during such scenarios. With the availability of wind power forecasts, the damage done to a battery can be assessed in terms of number of successive state changes from charging to discharging. State change q_i^+, $i \in \mathbb{N}$ can be defined as an event in a battery system which corresponds to a charging-discharging instance.

When a battery undergoes change in its state from charging to discharging, a state change is counted as 1. A charging event can be assigned discrete values of either 1 or 0 and vice-versa for discharging event. A simple algorithm for determining the number of state changes with the availability of measured and predicted wind powers is illustrated in Figure 2. The actual and forecasted wind powers are compared and a charging event is identified if the actual wind power exceeds the forecasted one. For a BESS, the discharging and charging powers are defined as

$$p_{\text{batt}} = \begin{cases} p_{\text{ch}} = P_w - \hat{P}_w > 0 \\ p_{\text{dis}} = \hat{P}_w - P_w > 0, \end{cases} \tag{6}$$

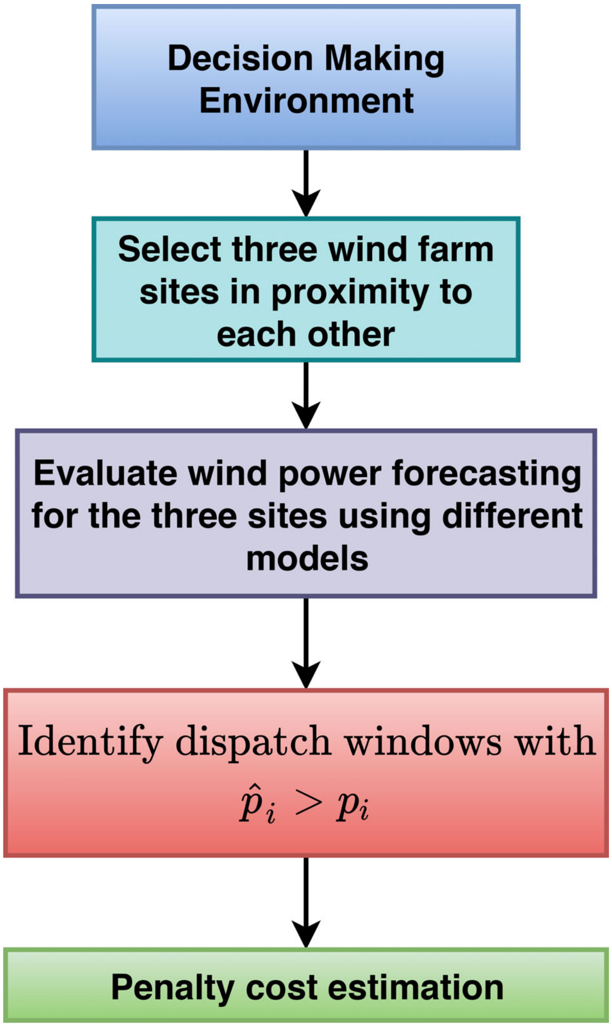

FIGURE 2 Battery state change algorithm based on charging and discharging powers

where P_w and $\hat{P}_w$ are the actual and predicted wind powers. In this regard, the number of consecutive battery state changes (charging to discharging) are associated with error in wind forecast. Consider standard error in terms of $\left(P_w - \hat{P}_w\right)$, an increase in this error leads to continuous battery charging. Thus, minimizing this error can certainly prevent violation of SoC limits. With the help of machine learning forecasting techniques, it is possible to minimize the aforementioned error and thus one can estimate the condition of battery charging-discharging profile. The charging of BESS is only done if the battery is in a stage where it can accept power, that is, when the SoC limits are not violated. Further, BESS is allowed to discharge if the actual wind power falls short of the forecasted one. This charging and discharging instance of BESS is assigned 1 and 0 respectively. A counter NchE(i) = 1 indicating charging instance and NchE(i) = 0 indicating discharging instance, is evaluated. Further, for N samples, the difference for each consecutive value of NchE is determined and a state change $q_i^+ = 1$ is assigned if the difference is 1, else 0 is assigned.

Higher number of state changes that arise from poor wind forecasting algorithm can degrade the battery life and necessitates frequent maintenance and the cost incurred can be as high as 11 000 USD/year.[31] Imposing a penalty cost for consecutive state changes will induce the operator to use an efficient scheme. This penalty cost is expressed as

$$N_{\mathrm{sc}} = \left\langle q_1^+, q_1^+ + q_2^+, \dots, \sum_{i=1}^{N} q_i^+ \right\rangle, \; F_{\mathrm{cost}} = \kappa N_{\mathrm{sc}}, \tag{7}$$

where κ is the penalty cost paid per consecutive state change in a battery system, N_{sc} represents the number of state changes, and q_i^+ is a state change index, and cost F_{cost} is a metric to evaluate the battery health in terms of number of successive state changes. The cost corresponding to an efficient forecasting method is likely to benefit the operator.

3 | TWIN SUPPORT VECTOR REGRESSION

For an integrated solution, we achieve Penalty cost minimization via wind power forecasting by adopting a twin support vector regression (TSVR) method and compare with ε-TSVR, random forest, and gradient boosted machines. We use a wavelet transform based signal processing to decompose the wind power obtained from wind speed data into low and high-frequency components. The forecasting model's inputs are the approximate and detail signals, and wind power is taken as output.[22]

Xinjun[18] paved a way for the solution of regression based problems by formulating a twin-hyperplane based support vector machines known as twin support vector regression. This technique computes the nonparallel hyperplanes around the data points by solving two quadratic programming problems (QPPs). Similar to classical SVR, TSVR evaluates two ε-insensitive functions which are up-bound and down-bound regressors. For training data $(x_1,y_1),(x_2,y_2),\dots,(x_n,y_n) \subset X \times \mathrm{R}$, where X represents the input feature space with dimension R^n, consider $Y = (y_1, y_2, \dots, y_i)$ as target output, $i = 1,2,\dots,n$ and $y_i \in \mathrm{R}$. The mathematical formulation of TSVR is

$$\begin{aligned}
&\min \frac{1}{2}\sum_{i=1}^{n}(y_i - e\varepsilon_1 - \psi_{1i})^T(y_i - e\varepsilon_1 - \psi_{1i}) + C_1 e^T \sum_{i=1}^{n}\xi_i, \\
&\qquad\qquad \mathbf{s.t.}\;\; y_i - \psi_{1i} \geq e\varepsilon_1 - \xi_i, \\
&\min \frac{1}{2}\sum_{i=1}^{n}(y_i - e\varepsilon_2 - \psi_{2i})^T(y_i - e\varepsilon_2 - \psi_{2i}) + C_2 e^T \sum_{i=1}^{n}\eta_i, \\
&\qquad\qquad \mathbf{s.t.}\;\; \psi_{2i} - y_i \geq e\varepsilon_2 - \eta_i,
\end{aligned}$$

where $\psi_{1i} = x_i w_1 + e b_1, \psi_{2i} = x_i w_1 + e b_2$, $C_1, C_2 > 0$ and $\varepsilon_1, \varepsilon_2 \geq 0$ are the TSVR hyperparameters and ξ_i, η_i denote the slack variables acting as soft margin to the error ε. The formulation of dual of the TSVR problem can be expressed in terms of Lagrangian operator given as

$$\begin{aligned} L(w_1,b_1,\varepsilon_i,\alpha_i,\beta_i) = \frac{1}{2}\sum_{i=1}^{n}(y_i - e\varepsilon_1 - (x_i w_1 + eb_1))^T \\ (y_i - e\varepsilon_1 - (x_i w_1 + eb_1)) \\ + C_1 e^T \sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i(y_i - e\varepsilon_1 - (x_i w_1 + eb_1)) - \sum_{i=1}^{n}\beta_i\xi, \end{aligned} \tag{8}$$

where α_i,β_i for $(i=1,2,...,n)$ are the Lagrangian multipliers. The KKT conditions can be evaluated as follows

$$\begin{cases} \dfrac{\partial L}{\partial w_1} = 0 \Rightarrow -X^T(Y - Xw_1 - eb_2 - e\varepsilon_1) + X^T\alpha = 0 \\ \dfrac{\partial L}{\partial b_1} = 0 \Rightarrow -e^T(Y - Xw_1 - e\varepsilon_1 - eb_2) + e^T\alpha = 0 \\ \dfrac{\partial L}{\partial \xi} = 0 \Rightarrow C_1 e^T - \alpha - \beta = 0 \\ \dfrac{\partial L}{\partial \alpha} = 0 \Rightarrow Y - (Xw_1 + eb_1) \geq e\varepsilon - \xi,\ \xi \geq 0, \end{cases}$$

For the TSVR optimization problem, the equality constraints are given as

$$\alpha^T(Y - (Xw_1 + eb_1) - e\varepsilon_1 + \xi) = 0,\ \alpha = 0,\ \beta^T\xi = 0 \tag{9}$$

where $\alpha \in [0, C_1 e]$ for $\beta \geq 0$, and (9) is unified and written as

$$-\begin{bmatrix} X^T \\ e^T \end{bmatrix}\left((Y - e\varepsilon_1) - [X\ \ e]\begin{bmatrix} w_1 \\ b_1 \end{bmatrix}\right) + \begin{bmatrix} X^T \\ e^T \end{bmatrix}\alpha = 0. \tag{10}$$

Let us define

$$Q = [X \quad e],\ t = Y - e\varepsilon_1,\ u_1 = \left[w_1^T \quad b_1\right]^T, \tag{11}$$

$$\begin{aligned} -Q^T t + Q^T Q u_1 + Q^T\alpha = 0, \\ u_1 = \left(Q^T Q\right)^{-1} Q^T(t - \alpha) \end{aligned} \tag{12}$$

where $Q = [X\,e]$ and $t = Y - e\varepsilon_1$. The matrix $\left(Q^T Q\right)$ is a positive semi-definite one with positive eigen values. This matrix is often an "ill-conditioned" matrix, where calculating its inverse can stir up computation errors and as a solution, a small regularization term σI, of order 10^{-7}, and I of appropriate dimensions is added to it. The dual corresponding to (8) can be simplified as

$$\begin{aligned} \max \quad -\frac{1}{2}\alpha^T Q\left(Q^T Q\right)^{-1} Q^T\alpha + t^T Q\left(Q^T Q\right)^{-1} Q^T\alpha - t^T\alpha \\ \textbf{s.t.} \quad \alpha \in [0, C_1] \end{aligned} \tag{13}$$

$$\begin{aligned} \max \quad -\frac{1}{2}\gamma^T Q\left(Q^T Q\right)^{-1} Q^T\gamma + m^T Q\left(Q^T Q\right)^{-1} Q^T\gamma - m^T\gamma \\ \textbf{s.t.} \quad \gamma \in [0, C_2], \end{aligned} \tag{14}$$

where $m = Y + e\varepsilon_2$ and $u_2 = \left(Q^T Q\right)^{-1} Q^T(m - \gamma)$. Equations (13) and (14) refer to the dual of original convex optimization problem. The final predictions for new data samples can be expressed in form of a mean regressor

$$f_{\text{TSVR}}(x) = \frac{1}{2}\left((w_1 + w_2)^T x + (b_1 + b_2)\right). \tag{15}$$

Next, we discuss the simulation results obtained under the decision-making environment for a hybrid wind-battery system for penalty cost minimization.

4 | RESULTS AND DISCUSSIONS

For penalty cost estimation by forecasting the wind power, the performance of Twin support vector regression (TSVR), as discussed in the previous section, is compared with ε-Twin support vector regression (ε-TSVR), random forest (RFR), and gradient boosted machines (GBM). For validation with real-time wind data, three offshore wind farm sites namely, Gemini (Netherlands), Veja Mate (Germany), and Walney (UK) located on the western coast of Denmark and Germany, capable of generating bulk power owing to high wind resource, are chosen. Further, three onshore wind farm sites namely, Clyde (Scotland), McCain Foods (UK), and Nygårdsfjellet (Norway) are also taken into consideration. The wind speed data for all the six wind farm sites is collected from MERRA-2[32] for the month of March 2019 with a 10 minute sampling interval. The wind turbine diameter is taken as 120 m. Out of 4320 samples collected for wind speed data, 3000 samples are used for training and remaining for testing. Further, training phase is treated with 10-fold cross-validation. Further, the forecast horizon is one-day ahead in blocks of 10-minutes. The hyper-parameters for TSVR, ε-TSVR are tuned in set 2^i, where $i = -9, -8, \ldots, 10$. Further, for random forest and gradient boosted machines, the number of trees are tuned from 1 to 10000.

In Table 1, the root mean square error for wind speed prediction is depicted. This analysis is carried out to validate that a similar pattern is observed in penalty cost. TSVR based prediction technique results in superior wind speed forecasting followed by ε-TSVR, RFR, and GBM. The basis of wind power forecasting is a hybrid model using wavelet transform and an ML algorithm. Penalty cost assessment with dispatch window of 10 minutes is described in Figure 1. The values of $\beta_w = 0.9$, $\zeta_s = 0.75$, $\alpha_z = 0.4$, and $\delta_x = 0.8$ in \$/kW are chosen so as to implement an accurate wind forecasting method. Table 2 highlights the penalty cost incurred to each wind farm site's operator for a given dispatch horizon. Considering the penalty costs for offshore sites, we find that TSVR based forecasting method results in minimum cost among the four regressors as indicated by bold numerals. For wind farm site Gemini, TSVR results in 85.74%, 92.40%, and 93.25% saving in penalty cost compared to ε-TSVR, RFR, and GBM methods, respectively. Similarly, for the site Veja Mate, TSVR yields 92.05%, 96.41%, and 95.83% saving when compared to other models as listed in Table 2.

The accuracy in wind power forecasts for TSVR and ε-TSVR are achieved by tuning the hyper-parameters like bandwidth (σ) of radial basis kernel function and regularization constant (C_1, C_2).[33] Forecast quality, which is often assessed by the coefficient of determination (R^2), varies with hyper-parameters.[22] Similarly, with RFR and GBM models, hyper-parameter tuning in several trees and learning rate affects the prediction. The mean squared loss, which is a measure of accuracy, decreases with an increase in trees number. Penalty costs are subjected to a change with a hyper-parameter variation for TSVR, ε-TSVR, RFR, and GBM model. As discussed, by knowing the number of state changes apriori, we calculate the cost incurred due to consecutive state changes. The value of κ is taken as \$15, and Table 3 depicts the penalty cost paid for consecutive state changes for BESS with different forecasting methods. The parameter κ demonstrates the penalty cost imposed on the farm operator for consecutive battery state changes. A consecutive change in battery state, that is, charging to discharging leads to degradation in its life. In this manuscript, we have assumed a value of

TABLE 1 Performance metric using selected ML methods

Root mean square error (RMSE)			
Method	**Gemini**	**Veja Mate**	**Walney**
TSVR	**1.1103**	**1.5805**	**1.1213**
ε-TSVR	2.2908	3.1788	2.6715
RFR	3.4254	4.2732	4.0887
GBM	4.0778	5.1316	4.6597

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

TABLE 2 Penalty cost (J^*) per 1000 USD with four forecasting models

Offshore wind farm sites			
Method	**Gemini**	**Veja Mate**	**Walney**
TSVR	**212.62**	**109.65**	**77.07**
ε-TSVR	1491.70	1380.70	716.37
RFR	2798.9	3057.90	1532.60
GBM	3151.40	2631.9	1523.10
Persistence	3312.6	2815.3	1735.80
Onshore wind farm sites			
Method	**Clyde**	**McCain Foods**	**Nygå rdsfjellet**
TSVR	**31.72**	**39.40**	**86.35**
ε-TSVR	265.92	751.45	530.24
RFR	5519.60	11 627	2007.20
GBM	6184.40	10 155	1472.1
Persistence	7132.1	12 158.3	2165.16

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

TABLE 3 Penalty cost (in $) for successive battery state changes

Offshore wind farm sites			
Method	**Gemini**	**Veja Mate**	**Walney**
TSVR	**1245**	**1965**	**1380**
ε-TSVR	1995	2940	2655
RFR	2940	3630	3585
GBM	4995	6480	5610
Onshore wind farm sites			
Method	**Clyde**	**McCain Foods**	**Nygårdsfjellet**
TSVR	**150**	**450**	**570**
ε-TSVR	1005	780	2355
RFR	2565	1095	1965
GBM	3495	2925	3570

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

USD 15/kW which is 6% of the capital cost in the range of (200-250 USD/kW). The range of the penalty cost is kept on the higher side in order to encourage the farm operator to adopt accurate forecasting technique.

From Figure 3, we observe the quantitative aspect of a battery system for various wind farm sites. From Table 3, we find that for all the wind farm sites, TSVR based method yields minimum state changes and thus results in minimum penalty cost followed by ε-TSVR, RFR, and GBM model. The variations of penalty cost T_2 with battery threshold power (P_b^{th}) are illustrated in Figure 4.

The cost incurred in strategy T_2 allows a wind farm operator to conserve the battery health in multiple charging-discharging instances. The cost approaches a minimum value for a given threshold battery power expressed in maximum discharging power. The superiority of TSVR over other methods is reflected in the penalty costs calculated in Tables 2 and 3, and Figure 4. From Figure 4, we can observe that cost is minimum for $P_b^{\text{th}}/\max(P_{\text{dis}}) = 0.8$ and increases thereafter. This threshold battery power where the cost incurred T_2 is minimum, is called optimal threshold battery power $P_{b,\text{opt}}^{\text{th}}$. This cost is prone to increase with a poor forecasting method. Therefore, wind farm operators are always encouraged to use an accurate forecasting method. The point $P_{b,\text{opt}}^{\text{th}} = 0.8$ is zoomed and the same is illustrated in Figure 4 for different forecasting methods. Results reveal the minimum cost corresponding to optimal threshold battery

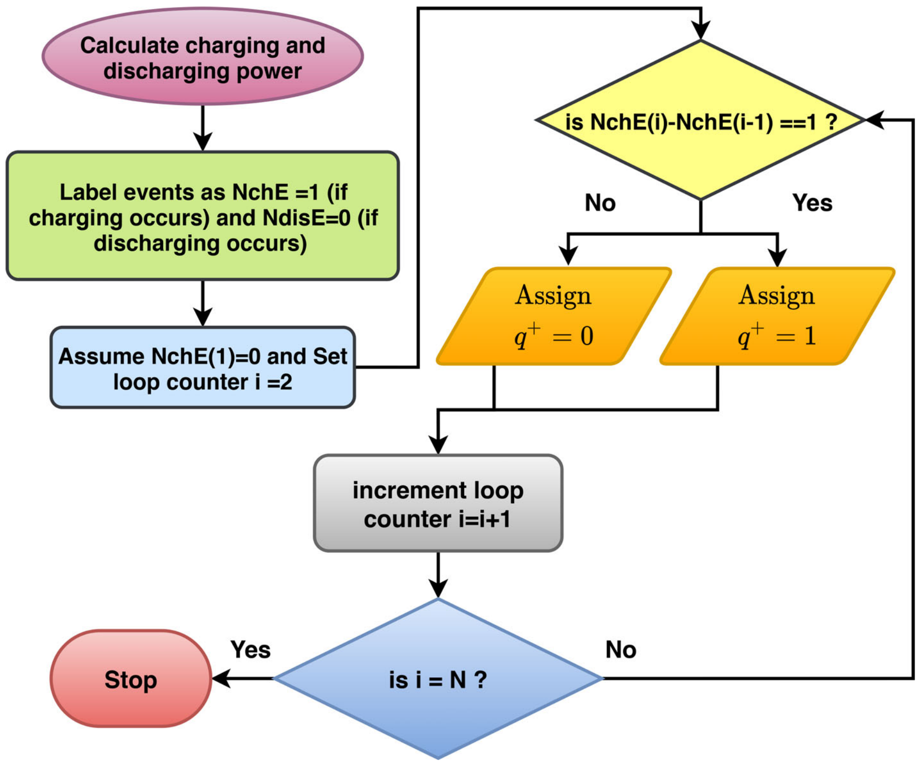

FIGURE 3 Number of distinct battery state changes for different wind farm sites under different forecasting models

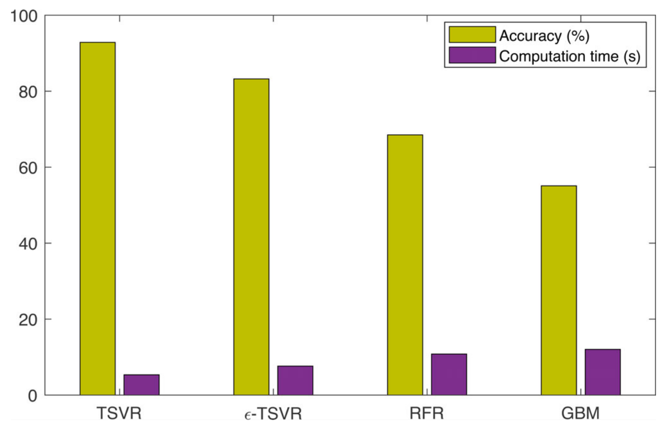

FIGURE 4 Penalty cost T_2 variation with threshold battery power

power for TSVR. The overall scenario of arriving at an integrated solution is based on finding the best possible regressor capable of minimizing the penalties imposed on the operator. Cost minimization is an important aspect that must be evaluated when dealing in electricity markets where trading is done on a minute to minute scale. Providing short-term wind forecasts to TSO benefits the wind farm operators especially when the need for reserve power capacity can escalate operating costs. To evaluate the scenario for a hybrid wind farm, two cases: (a) static and (b) dynamic are assessed. A static scenario is defined as one where the optimal threshold battery power ($P^{\text{th}}_{b,\text{opt}} = 0.8$) as discussed in penalty cost T_2, is available to wind farm operator to pay the respective penalty. On the other hand, a dynamic case for a wind farm operator is when the power available in BESS is not the optimal threshold battery power. In such a case, the farm operator has no option but to operate at a nonoptimal battery power which results in a higher penalty cost as illustrated in Figure 4. A common cost metric for each forecasting method needs to be evaluated in the static and dynamic scenario. Let us define V_1, V_2, V_3 as the three individual costs as J^*, T_2, and F_{cost} depicted in Table 2, Figure 4 and Table 3 respectively, and h_1, h_2, h_3 as the weights associated with these costs. A common cost metric is given as

$$CC_j = h_1 V_1 + h_2 V_2 + h_3 V_3, \tag{16}$$

where CC_j is the cost metric for the jth forecasting method. The weights h_1, h_2 and h_3 attain a value of 1 for static case but weight h_2 associated with cost V_2 needs to be tuned and lies in the range $(0,1]$. The static case where the weights h_1, h_2 and h_3 are assigned a value of 1 result in a common cost metric CC depicted in Table 4. It is observed that, for the static scenario, TSVR based forecasting method yields in minimum common cost for all the datasets. The increase in common cost metric as we change the forecasting method is associated with the accuracy of the method.

Thus, when a wind farm operator has the possibility to operate BESS at optimal threshold point, minimum cost is incurred and the weight assigned to cost V_2 is 1. For the dynamic scenario case, the weight h_2 can take values between 0 and 1. In a nonoptimal battery threshold scenario, the wind farm operator will attempt to use a better forecasting scheme which will ultimately reduce the overall cost. When the threshold battery power is less than the optimal threshold battery power ($P^{th}_{b,\mathrm{opt}} = 0.8$), the wind farm operator is likely to incur a higher operating cost as observed from the Figure 4. Thus the value of weight h_2 corresponding to this cost V_2 has to be assigned higher value to penalize the farm operator irrespective of the forecasting method used. This value of h_2 is tunable and may attain lesser values when a superior battery technology is used. On the other hand, if the threshold battery power is more than the optimal threshold battery power ($P^{th}_{b,\mathrm{opt}} = 0.8$), the farm operator has a better chance to incur lesser overall cost due to proximity to the optimal threshold power. Thus, in this case, the weight h_2 is assigned a value less than or equal to 0.5. The weight h_2 is expressed as

$$h_2 = \begin{cases} \leq 0.5, & \text{if } P^{th}_b > P^{th}_{b,\mathrm{opt}} \\ 0.5, & \text{if } P^{th}_b < P^{th}_{b,\mathrm{opt}} \end{cases}. \tag{17}$$

Consider a scenario where the weight h_2 attains a value of 0.75 and 0.4, the common cost metric CC corresponding to this scenario is depicted in Table 5.

As observed from these tables, the scenario with $h_2 = 0.4$, yields a lower overall cost than $h_2 = 0.75$. That is, when the threshold battery power available is near the optimal point, the cost incurred is less. It also strengthens the fact that a superior battery technology with a higher power density and discharge efficiency will allow the wind farm operator to operate near optimal battery threshold power. The superior battery technology will also be beneficial in providing sustainable solutions to the wind farm operators in terms of the sudden changes like that of wind speed ramp events. Surplus power from large power reversal caused due to ramp events can be stored in battery storage. In Table 6, the common cost metric is evaluated for a wind speed dataset having a length of 3 months. The wind speed data for sites is collected from May 2019 to July 2019. With the wind speed data from May 2019 being used for training and data corresponding to June to July 2019, that is a total of 8784 samples being used for the testing phase. It is observed that

TABLE 4 Common cost *CC* (in \$) for static scenario

Offshore wind farm sites			
Method	**Gemini**	**Veja Mate**	**Walney**
TSVR	**1465.60**	**2086.26**	**1480.13**
ε-TSVR	3551.98	4460.13	3526.46
RFR	5842.04	6960.68	5397.09
GBM	8286.77	9413.75	7470.01
Onshore wind farm sites			
Method	**Clyde**	**McCain Foods**	**Nygårdsfjellet**
TSVR	**187.10**	**492.98**	**660.79**
ε-TSVR	1303.95	1651.74	2962.77
RFR	8203.79	13 156.06	4092.30
GBM	9914.33	13 873.47	5165.44

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

TABLE 5 Common cost CC (in \$) for dynamic scenario

Method	Gemini	Veja Mate	Walney
$h_2 = 0.75$			
TSVR	**1463.60**	**2083.35**	**1474.37**
ε-TSVR	3535.66	4425.27	3487.69
RFR	5823.01	6892.48	5327.21
GBM	8251.68	9338.29	7385.78
$h_2 = 0.4$			
TSVR	**1460.81**	**2079.29**	**1466.29**
ε-TSVR	3512.81	4376.47	3433.40
RFR	5783.75	6797.01	5229.39
GBM	8202.55	9232.64	7267.86

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

TABLE 6 Common cost CC (in \$) for dynamic scenario for 3-month wind dataset

Method $h_2 = 0.75$	Gemini	Veja Mate	Walney
TSVR	**2463.51**	**3083.35**	**2474.37**
ε-TSVR	3415.36	4054.71	3853.29
RFR	6123.41	5992.18	4194.12
GBM	7528.28	8083.91	4956.83

Note: The bold values indicate the best regression model with respect to a specific wind farm site.

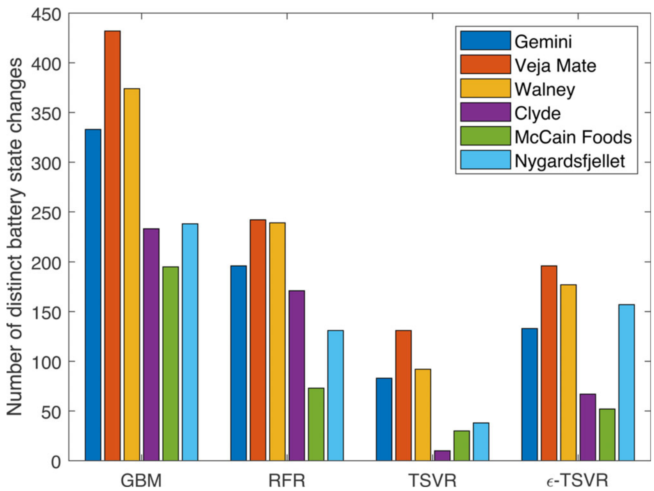

FIGURE 5 Comparison of accuracy and computation time for tested methods

TSVR based forecasting technique outperforms the rest in terms of the common cost metric. The magnitude of common cost increases compared to 1 month testing results as discussed in Table 5.

Furthermore, in terms of the best forecasting method for penalty cost minimization, the accuracy obtained from TSVR is the highest, followed by ε-TSVR, Random forest, and Gradient boosted machines. The high accuracy of the TSVR method is primarily due to the resultant regressor's ability to capture noise present in the wind speed dataset. The superior prediction performance of TSVR thus yields minimum penalty costs for the wind farms under consideration. Figure 5 illustrates the accuracy and the computation time of forecasting methods tested for penalty cost minimization. We observe that TSVR has the least computation time than ε-TSVR, RFR, and GBM. The provision in TSVR of solving two quadratic programming problems provides faster computation.

5 | CONCLUSIONS

This manuscript presents a machine learning-based integrated solution for penalty cost minimization of a hybrid wind-battery farm. We compare the wavelet-Twin support vector regression-based forecasting method with ε-TSVR, Random forest, and Gradient boosted machine. Maximizing the annual energy production (AEP) is a significant concern for a wind farm operator. For validation, we consider six wind farm sites with real-time wind speed data and find that the wavelet-Twin support vector regressor yields minimum cost followed by ε-TSVR, Random forest, and Gradient boosted machines. Furthermore, the impact of consecutive battery charging-discharging is also analyzed and penalized at a reasonable cost. Results reveal that wavelet-TSVR yields minimum penalty cost among the four tested methods. Wavelet-TSVR forecasting method results in an accuracy of 92.33%, followed by ε-TSVR with 82.45%, random forest with 67.89%, and Gradient boosted machines with 59.01%. In terms of computation speed, wavelet-TSVR is the fastest. Overall, for a hybrid wind-battery farm, an accurate wind forecasting scheme is an important parameter to materialize key events like imbalance in power generation and battery charging-discharging states.

CONFLICT OF INTEREST

Authors hereby confirm that there is no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/2050-7038.13010. [Correction added on 16 September 2021, after first online publication: Peer review history statement has been added.]

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Harsh S. Dhiman https://orcid.org/0000-0002-7394-7102

REFERENCES

1. Deng X, Lv T. Power system planning with increasing variable renewable energy: a review of optimization models. *J Clean Prod*. 2020; 246:118962. https://doi.org/10.1016/j.jclepro.2019.118962
2. Dhiman HS, Deb D, Muyeen SM, Kamwa I. Wind turbine gearbox anomaly detection based on adaptive threshold and twin support vector machines. *IEEE Trans Energy Convers*. 2021;1-1. https://doi.org/10.1109/tec.2021.3075897
3. Hau E, Renouard VH. *Wind Turbines: Fundamentals, Technologies, Application, Economics*. Berlin: Springer; 2006.
4. Kragh K, Hansen M, Mikkelsen T. *Improving Yaw Alignment Using Spinner Based LIDAR*. Orlando, FL: American Institute of Aeronautics and Astronautics; 2011.
5. Matevosyan J, Soder L. Minimization of imbalance cost trading wind power on the short-term power market. *IEEE Trans Power Syst*. 2006;21(3):1396-1404. https://doi.org/10.1109/tpwrs.2006.879276
6. Basit A, Hansen AD, Altin M, Sørensen PE, Gamst M. Compensating active power imbalances in power system with large-scale wind power penetration. *Journal of Modern Power Systems and Clean Energy*. 2015;4(2):229-237. https://doi.org/10.1007/s40565-015-0135-x
7. Egido I, Fernandez-Bernal F, Rouco L, Porras E, Saiz-Chicharro A. Modeling of thermal generating units for auto-matic generation control purposes. *IEEE Transactions on Control Systems Technology*. 2004;12(1):205-210. https://doi.org/10.1109/tcst.2003.821959
8. Liu H, Chen C, Lv X, Wu X, Liu M. Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods. *Energ Conver Manage*. 2019;195:328-345. https://doi.org/10.1016/j.enconman.2019.05.020
9. Deb D, Ambekar A, Sagi D. Method and system for hybrid wind power generation, 10066604 B2. US Patent; 2018.
10. Patel P, Shandilya A, Deb D. Optimized Hybrid Wind Power Generation with Forecasting Algorithms and Battery Life Considerations. In: IEEE; 2017.
11. Yuan X, Tan Q, Lei X, Yuan Y, Wu X. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine. *Energy*. 2017;129:122-137.
12. Du P, Wang J, Yang W, Niu T. A novel hybrid model for short-term wind power forecasting. *Appl Soft Comput*. 2019;80:93-106. https://doi.org/10.1016/j.asoc.2019.03.035
13. Wang J, Zhang W, Li Y, Wang J, Dang Z. Forecasting wind speed using empirical mode decomposition and Elman neural network. *Appl Soft Comput*. 2014;23:452-459.
14. Liu H, Chen C, Tian QH, Li FY. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renew Energy*. 2012;48:545-556.
15. Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. *Renew Energy*. 2016;94:629-636.

16. Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm—multi-objective whale optimization algorithm for wind speed forecasting. *Appl Energy*. 2017;208:344-360.
17. Suykens J, Vandewalle J. Least squares support vector machine classifiers. *Neural Processing Letters*. 1999;9(3):293-300.
18. Peng X. TSVR: an efficient twin support vector machine for regression. *Neural Netw*. 2010;23(3):365-372.
19. Shao YH, Zhang CH, Yang ZM, Jing L, Deng NY. An *s*-twin support vector machine for regression. *Neural Comput Appl*. 2012;23(1): 175-185.
20. Dhiman HS, Deb D. Wake management based life enhancement of battery energy storage system for hybrid wind farms. *Renew Sustain Energy Rev*. 2020;130:109912. https://doi.org/10.1016/j.rser.2020.109912
21. Dhiman HS, Deb D. Machine intelligent and deep learning techniques for large training data in short-term wind speed and ramp event forecasting. *Int Trans Electr Energy Syst*. 2021. https://doi.org/10.1002/2050-7038.12818
22. Dhiman HS, Deb D, Guerrero JM. Hybrid machine intelligent SVR variants for wind forecasting and ramp events. *Renew Sustain Energy Rev*. 2019;108:369-379. https://doi.org/10.1016/j.rser.2019.04.002
23. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. UK: Routledge; 2017.
24. Dhiman HS, Deb D, Muresan V, Unguresan ML. Multi-criteria decision making approach for hybrid operation of wind farms. *Symmetry*. 2019;11(5):675. https://doi.org/10.3390/sym11050675
25. Biswas PP, Arora P, Mallipeddi R, Suganthan PN, Panigrahi BK. Optimal placement and sizing of FACTS devices for optimal power flow in a wind power integrated electrical network. *Neural Comput Appl*. 2020;33:6753-6774. https://doi.org/10.1007/s00521-020-05453-x
26. Dhiman HS, Deb D. *Decision and Control in Hybrid Wind Farms*. Singapore: Springer; 2019.
27. Dhiman HS, Deb D, Balas VE. *Supervised Machine Learning in Wind Forecasting and Ramp Event Prediction*. Academic Press; 2019.
28. Grady S, Hussaini M, Abdullah M. Placement of wind turbines using genetic algorithms. *Renew Energy*. 2005;30(2):259-270. https://doi.org/10.1016/j.renene.2004.05.007
29. Parada L, Herrera C, Flores P, Parada V. Assessing the energy benefit of using a wind turbine micro-siting model. *Renew Energy*. 2018; 118:591-601. https://doi.org/10.1016/j.renene.2017.11.018
30. Abdelsalam AM, El-Shorbagy M. Optimization of wind turbines siting in a wind farm using genetic algorithm based local search. *Renew Energy*. 2018;123:748-755. https://doi.org/10.1016/j.renene.2018.02.083
31. Li X, Chalvatzis K, Stephanides P. Innovative Energy Islands: life-cycle cost-benefit analysis for battery energy storage. *Sustainability*. 2018;10(10):3371. https://doi.org/10.3390/su10103371
32. Soda-Pro. The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) Web service. http://www.soda-pro.com/web-services/meteo-data/merra. Accessed on November 5, 2021.
33. Dhiman HS, Anand P, Deb D. *Wavelet Transform and Variants of SVR with Application in Wind Forecasting*. Singapore: Springer; 2018: 501-511.

How to cite this article: Dhiman HS, Deb D, Muyeen SM, Abraham A. Machine intelligent forecasting based penalty cost minimization in hybrid wind-battery farms. *Int Trans Electr Energ Syst*. 2021;31(9):e13010. https://doi.org/10.1002/2050-7038.13010