Journal of Theoretical and Applied Information Technology
30th June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A HYBRID VARIATIONAL QUANTUM–DEEP LEARNING FRAMEWORK FOR ELECTRICAL LOAD FORECASTING IN SMART GRIDS

RAVURI DANIEL[1], BODE PRASAD[2], PASAM PRUDHVI KIRAN[3], KANDRAKUNTA CHINNAIAH[4], INDHUMATHI RAVICHANDRAN[*5], RAMANJANEYULU DAYINABOYINA[6], MANIDHEER BABU GORIKAPUDI[7], DORABABU SUDARSA[8]

[1]Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, India
[2]Department of Information Technology, Vignan's Institute of Information Technology (A), Visakhapatnam, India
[3]Department of Computer Science and Engineering (IoT), R.V.R. & J. C. College of Engineering (A), Guntur, India
[4]Department of CSE, MLR Institute of Technology, Dundigal, Hyderabad, India
[5*]Department of Electrical and Electronics Engineering, M. Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
[6]Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India
[7]Department of CSE-Artificial Intelligence and Machine Learning, Vignans Nirula Institute of Technology and Science for Women, Guntur, India
[8]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, India
E-mail: [1]danielravuri@gmail.com, [2]arjunprasad.bode@gmail.com, [3]pasamprudhvi@gmail.com, [4]chinna.nitc@gmail.com, [5*]indhumathirjs@gmail.com, [6]dramanjaneyulu@kluniversity.in, [7]manidheer.gorikapudi@gmail.com, [8]dorababu.sudarsa@gmail.com

ABSTRACT

In the smart grid, electrical load forecasting plays an important role in its operation, especially as renewable energy sources are increasingly integrated into the grid, increasing demand variability. Recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks can learn temporal patterns but are often ineffective at handling long-range dependencies, complex feature relationships, and fluctuations caused by renewable energy sources. In this study, a hybrid Variational Quantum–Deep Learning model for short-term electrical load forecasting is proposed. To enhance feature representation, sequence learning, and time-step weighting, the model incorporates a Variational Quantum Circuit (VQC), LSTM layers, and an attention mechanism. Temporal forecasting is preceded by preprocessing and encoding of historical load data and exogenous variables, such as weather, calendar, and renewable-energy indicators. The proposed model is compared with the baseline models, such as Multilayer Perceptron (MLP), RNN, LSTM, Convolutional Neural Network–LSTM (CNN-LSTM), and the coefficient of determination (R^2) using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The forecasting errors of the proposed Variational Quantum Algorithm–LSTM–Attention (VQA–LSTM–Attention) model are found to be the smallest, with RMSE and MAPE 22.64% and 25.64% lower than those of the standalone LSTM model, respectively. These results show that the accuracy of short-term load forecasting in uncertain smart grids can be enhanced by combining quantum-enhanced feature transformation with LSTM and attention-based learning. The paper proposes an applicable hybrid quantum–classical forecasting model for demand response, renewable integration, market bidding, and grid stability.

Keywords: *Electrical Load Forecasting; Variational Quantum Circuit (VQC); Long Short-Term Memory (LSTM); Quantum Machine Learning; Smart Grids.*

1. INTRODUCTION

Electrical load forecasting is one of the key components of the modern smart grid, as it can be used to schedule generation, plan reserves, implement demand response, implement electricity pricing, bid for electricity, and provide real-time security for the grid. Forecasting accuracy is vital to the efficient and reliable operation of power systems in an increasingly digital, distributed, and interactive world. The smart grid now includes smart meters, distributed energy resources, electric vehicles, storage systems, prosumers, and dynamic tariffs, which generate more complex demands than those in a traditional power grid. Inaccurate short-term load forecasting can cause over-scheduling, higher balancing costs, reliability risks, and renewable curtailment, while accurate forecasting can enable unit commitment, economic dispatch, and congestion management, peak shaving, and market participation [1]–[3]. This has become an even more important task due to the surge in electricity demand. In 2024, global electricity demand rose by 4.3% from the previous year, just under twice the average annual increase over the previous decade, contributing around 1,080 TWh [1].

The growing mix of renewable energy resources has made load forecasting more complex. The intermittency of solar and wind power, their weather dependence, and their lack of dispatchability introduce uncertainty into energy supply and net load. Renewable electricity generation is projected to grow from 9900 TWh in 2024 to 16200 TWh in 2030, accounting for 32% to 43% of global electricity generation, with the share of variable renewables increasing from 10% to 19% and almost doubling by 2030 [2]. In this context, load forecasting is not just a problem of estimating the load, but rather a high-dimensional spatio-temporal prediction problem that is affected by weather conditions, calendar effects, consumer behavior, market signals, PV production, wind variability, and the interactions with storage [4] and [5]. Initial methods of forecasting electrical loads included statistical and classical machine-learning models, such as autoregressive integrated moving-average models, support vector machines, regression trees, random forests, and multilayer perceptrons. While these approaches were effective in stable-demand scenarios, they were labor-intensive to engineer features and did not always capture long-term, nonlinear temporal relationships. Deep learning models enhanced forecasting by facilitating feature extraction and sequential modeling based on data. The introduction of Recurrent Neural Networks (RNNs) enabled temporal memory, and Long Short-Term Memory (LSTM) networks demonstrated better solutions to the vanishing gradient problem and improved long-horizon sequence learning. LSTM-based recurrent networks have proven effective for short-term residential load forecasting, as shown by Kong et al. [6]. Subsequent research focused on highlighting key time steps and inputs [7], and decomposition methods such as VMD-LSTM and DWT-LSTM improved forecasting accuracy in the presence of noisy, nonstationary loads [8] and [9]. Current deep learning methods have yet to overcome significant shortcomings. The RNN, LSTM, GRU, CNN-LSTM, and transformer-based models can be very data-hungry, labor-intensive, and time-consuming to train and tune. These requirements make the problem sensitive to overfitting when the data are sparse, noisy, and/or region-specific [3], [10]. Scalability is also challenging due to diverse consumer types, dispersed renewable energy production, and multiple time resolutions in real-world smart grids. Moreover, higher penetration levels of renewables, behavioral changes, electrification, tariff changes, and weather anomalies can lead to inaccurate models when demand patterns shift. More recent studies in Applied Energy have highlighted the dynamism of energy production systems and the deterioration of model performance as new data patterns emerge [10]. Such constraints highlight the need for upcoming forecasting models that can better represent nonlinear features, learn over time, and be robust to renewable-driven uncertainty.

New methods are needed to solve energy forecasting problems involving many complex features, and Quantum computing offers a promising solution. Alternative ways of processing high-dimensional information are quantum states, superposition, entanglement, and parameterized quantum circuits. While fault-tolerant quantum computers are not yet widely available, Variational Quantum Algorithms (VQAs) are suitable for noisy intermediate-scale quantum devices, as they involve parameterized quantum circuits coupled with classical optimizers [11]. In energy research, both quantum and hybrid quantum–classical models have been studied for applications such as demand response, energy optimization, power-flow analysis, and load forecasting [12]–[14]. But the basic quantum models are limited by the number of qubits, hardware noise, circuit trainability, and scalability. Thus, a hybrid quantum–classical approach is a viable solution, bridging the representational power of quantum circuits and the well-known temporal modeling power of deep learning. The study's research gap concerns the lack of guidance on incorporating quantum feature representations into state-of-the-art temporal deep learning architectures for short-term electrical load forecasting. Classic deep-learning models can handle sequential data, but they might struggle with complex feature interactions, overfitting, and generalizing in the face of uncertainty in renewable energy. However, quantum-only models are not yet sufficiently developed for large-scale practical use. To fill this gap, this paper proposes a hybrid Variational Quantum–Deep Learning framework combining a Variational Quantum Circuit (VQC) with LSTM layers and an attention mechanism. The proposed architecture involves the VQC as a quantum feature-transformation layer, the LSTM for modeling sequential demand patterns, and the attention mechanism to focus on the most relevant time steps and demand-related features. The aim of this research work is to design and test a hybrid VQC–LSTM–Attention model for short-term EWF in a smart-grid environment.

Although electrical load forecasting has made a dramatic improvement, the current forecasting models are still facing a huge challenge in the smart-grid environment. There are many highly nonlinear relationships between these factors (load demand, weather, renewable energy generation, and consumer behavior) that are not well represented by traditional statistical and machine-learning methods. Deep learning models like LSTM, CNN-LSTM, and attention-based models have enhanced forecasting accuracy but often require extensive data training, high computational power, and are prone to overfitting and reduced performance in scenarios with varying demand patterns. In addition, the growing share of renewables brings new uncertainty and variability which traditional forecasting models cannot properly capture. Although quantum machine learning has recently become a promising paradigm for learning complex feature relationships using high dimensional quantum representations, the current studies on quantum forecasting are still limited and are rarely combined with powerful temporal learning architectures. Thus, a hybrid forecasting model which integrates quantum enhanced feature representation and deep temporal learning mechanisms is needed to enhance forecasting accuracy, robustness and adaptability in renewable-rich smart-grid environments. A proposed Variational Quantum Circuit–LSTM–Attention framework is designed to overcome these drawbacks and offer a more efficient way to solve the challenges facing intelligent energy management systems of the future.

This study is significant because it can contribute to the development of next-generation intelligent forecasting systems for smart grids based on renewable energy. The proposed framework combines Variational Quantum Circuit (VQC) based feature transformation, Long Short-Term Memory (LSTM), and attention-based learning, providing a structured methodology for improving the short-term load prediction in the presence of uncertainty. A hybrid Variational Quantum–Deep Learning (VQ-DL) framework for short-term electrical load forecasting (ELF) in smart grid is proposed. A quantum feature representation model captures the quantum system's features, using an LSTM-based temporal model and an attention mechanism. The proposed model is compared with the classical model and deep learning-based models such as Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), LSTM, and Hybrid LSTM. The model is evaluated for forecasting performance under renewable-influenced load conditions with Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

This paper reviews the three main methods of load forecasting: Classical, Deep-learning, and Quantum machine-learning, in Section 2. In Section 3, the proposed hybrid Variational Quantum–Deep Learning (VQ-DL) methodology, including the design of the quantum circuit and the model architecture, is introduced. The results are discussed and compared with existing baselines in Section 4. The paper concludes in Section 5.

2. RELATED WORK

The field of electrical load forecasting has moved from statistical to shallow machine-learning-based methods, to deep learning, and to novel quantum–classical methods. This change reflects the increasing complexity of smart-grid systems, in which electricity consumption is influenced by non-linear demand, renewable energy supply, electric vehicles, storage systems, and dynamic electricity pricing [3], [15], [16]. For stable, structured datasets, initial methods based on autoregressive models, support vector regression, random forests, regression-based models, and multilayer perceptrons produced useful baselines [14] [15]. However, these approaches relied on manually designed features and were not very good at detecting nonlinear temporal dependencies. However, as the amount of data that the smart grid is expected to collect became more complex and diverse, the effectiveness of the smart grid was reduced because the smart grid wasn't able to accurately model: weather-induced demand, rooftop photovoltaic (PV) generation, charging of electric vehicles (EVs), tariff changes, and consumer habits [15, 16].

Some of the above limitations have been overcome by using deep learning models that learn features directly from sequential data. Long Short-Term Memory (LSTM) networks have become popular recently because they can overcome the vanishing gradient problem in traditional Recurrent Neural Networks (RNNs) and retain information for long periods, thereby enabling longer-term forecasting. Kong et al. argued that LSTM-based recurrent networks could be used to perform short-term residential load forecasting using smart-meter data [6]. Similarly, Fekri et al. highlighted the importance of adaptive recurrent models that can adapt to changing consumption behavior over time [17]. But LSTM models may require massive amounts of data, intensive computational resources, and careful hyperparameter tuning. They can also suffer from overfitting when the training data is scarce, noisy, or geographically localized [3, 15]. Some methods have improved LSTM performance through feature selection [8], genetic algorithm optimization [9], and decomposition-based preprocessing [18], but these enhancements make the LSTM architecture more complex and may reduce deployment efficiency. To overcome these drawbacks, hybrid deep learning models are developed by combining multiple learning mechanisms. While Gated Recurrent Units (GRUs) are simpler and faster than LSTMs, they might not perform as well as LSTMs when dealing with datasets that exhibit complex long-term memory, such as the output from houses with fluctuating demand or renewable energy sources with varying outputs [6], [7]. The Convolutional Neural Network-LSTM (CNN-LSTM) models combine local feature extraction and sequence learning, with the CNN layers detecting peaks and short-term demand variations, and the LSTM layers learning temporal dependencies. Attention-based models, which give more weight to time steps in the history that are more influential and better explain variables, improve forecasting [19], [20]. Self-attention extends this idea to transformer architectures, which can process long-range dependencies without recurrent processing [21, 22]. However, these models require substantial data and can be resource-intensive, limiting their applicability to smaller data sets, on-the-fly applications, and resource-constrained grid environments.

The analyzed studies show that deep learning models have improved forecasting accuracy, but unresolved issues remain. Many models are sensitive to data quality and availability, especially when missing values, noise, abnormal values, or region-specific patterns are present in the set of smart-meter records [3], [15]. These models are also limited in transferability because a model trained in one household, feeder, or city may not transfer to another context with different weather, occupancy, socioeconomic behavior, tariff structures, and renewable penetration [16], [23]. In addition, concept drift caused by weather extremes, electrification, shifts in consumer behavior, new appliances, tariff changes, or more renewable penetration [10, 17] can negatively impact forecasting. These restrictions point to the fact that the main problem is not only to minimize prediction error but also to increase the robustness and the generalization of the prediction at unknown grids. In recent years, quantum machine learning has been recognized as a viable approach to improving feature representation in complex forecasting problems. Qubits, superposition, entanglement, and quantum transformations of states are used in quantum computing to process information in ways different from classical computing [11,24, 25]. Since large-scale fault-tolerant quantum computers are not widely accessible, research has focused on Variational Quantum Algorithms (VQAs) that incorporate parameterized quantum circuits and classical optimizers [11, 26]. Variational Quantum Circuits (VQCs) can be used as trainable feature-mapping layers. The concept of quantum circuit learning has been proposed for nonlinear function approximation [27], and the use of quantum-enhanced feature spaces has been shown to facilitate supervised learning by transforming data into a high-dimensional Hilbert space [28]. The features are relevant to load forecasting due to the nonlinearity and dimensionality of electricity demand, as well as the impact of uncertain external variables such as prices, weather, and renewable electricity production.

As far as energy applications are concerned, the most explored ones are optimization, scheduling, demand response, power-flow analysis, and control [12], [29], [30]. Variational Quantum Circuits (VQC)-based hybrid approaches have been shown to be promising for deployment in demand response applications [12], and quantum artificial intelligence (AI) has been investigated for electrical load forecasting [13]. However, current research on quantum energy remains relatively limited. Most studies focus on optimization, and the few models for quantum forecasting are rarely contrasted with powerful deep learning benchmarks such as CNN-LSTM, Attention-LSTM, transformer models, or decomposition-based LSTM models [7]–[9] and

[13]. Other challenges for purely quantum models include limited qubit availability, circuit noise, barren plateaus, simulation overhead, and scalability challenges [11, 26]. There is a clear gap in the literature between the classical deep-learning forecasting approach and the new machine-learning techniques based on quantum computers. Although simple classical and shallow models exist, they can't be used for nonlinear temporal modeling. Although deep-learning models excel at sequence learning, they are prone to overfitting, poor transferability, high computational demands, and degraded performance in the face of uncertainty due to renewables. While there is some hope in utilizing quantum machine learning for feature representation, the quantum-only methods are not yet sufficiently developed for deployment in large-scale practical applications. The current study aims to fill this gap and introduces a hybrid Variational Quantum–Deep Learning approach to short-term electrical load forecasting. The model proposed here consists of 3 main components: Variational Quantum Circuit (VQC) for the nonlinear quantum feature transformation, Long Short Term Memory Layer (LSTM) for the modeling of the sequential nature of the demand patterns, and Attention Mechanism (Attn) for the selection of the most relevant time steps and input features. This integration addresses some of the limitations identified in previous studies by combining the quantum-enhanced feature representation with the temporal deep-learning components of smart-grid forecasting.

3. METHODOLOGY

This study introduces a hybrid Variational Quantum–Deep Learning approach for short-term electrical load forecasting in smart-grid environments. The proposed method combines a Variational Quantum Circuit (VQC) with a classical Long Short-Term Memory (LSTM) network and an attention mechanism to leverage quantum feature representations and classical temporal sequence learning. The goal of this hybrid design is to improve the model's ability to describe the temporal, uncertainty, and nonlinear aspects of electrical load demand. The research is an experimental, model-development design. The forecasting framework is driven by historical electrical load data and by relevant exogenous variables. The methodological process consists of seven main steps: data collection, data preprocessing, feature selection, quantum feature encoding, variational quantum transformation, classical temporal modeling by LSTM and attention, and end-to-end load prediction by a dense regression layer. The entire procedure is designed so that the forecasting model can be systematically replicated with the same input variables, preprocessing steps, model architecture, and evaluation. The overall architecture of the proposed hybrid VQC–LSTM model is shown in Figure 1. It starts with preprocessed smart-grid data whose quantum states are created by applying quantum feature encoding. These encoded states are then passed through the VQC layer, which produces quantum-enhanced features. The generated quantum features are then combined with classical input features and fed into the LSTM-attention module to learn from temporal features and to make the final prediction.

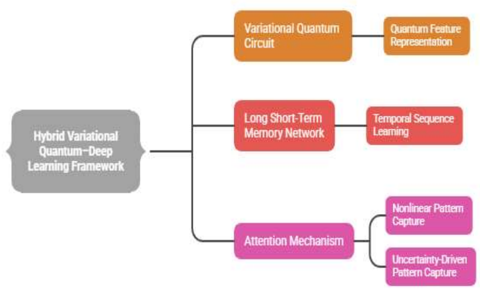

Figure 1. Hybrid VQC–LSTM architecture for load forecasting.

3.1 Hybrid Quantum–Classical Workflow

The proposed research design is a hybrid quantum–classical forecasting approach. The classical part is responsible for preprocessing, sequence modeling, and regression-based prediction, while the quantum part handles nonlinear feature transformation. The design allows the model to combine the advantages of quantum feature mapping and temporal learning provided by deep neural networks. A detailed overview of the proposed hybrid quantum–classical forecasting framework is provided in Figure 2. The data inputs are historical electrical load observations and exogenous data such as weather, renewables generation indicators, and calendar information. These variables are chosen because they are affected by weather variations, renewable energy variability, weekday/weekend effects, seasonal effects, and time-varying demand.

Each observation in this dataset corresponds to a time step in a time-series forecasting task. The dependent variable is the predicted future electrical load, and the independent variables are historical load values and exogenous historical predictors. The forecasting problem is cast as a one-step-ahead prediction problem, in which the model predicts the load at time t+1 based on data up to time t.

Real data inputs can be missing, strange values, noise, and different feature scales. Hence, a pre-processing function is used prior to model training. The processed feature vector is represented by:

$$x_t' = P(x_t) \quad (1)$$

The raw input vector at time t is denoted as x_t, the processed feature vector at time t is denoted as x'_t and the preprocessing operator *P(.)* refers to the cleaning, normalization and feature selection process.

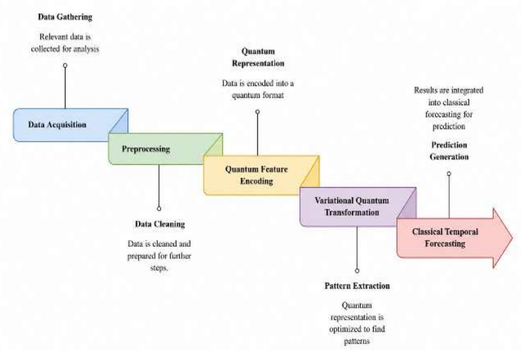

Figure 2. Workflow pipeline of the proposed framework.

Once the preprocessed classical features are selected, the selected classical features are encoded into the quantum states by the feature-encoding techniques, such as amplitude encoding or angle encoding. The step is the classical input data is mapped into a quantum Hilbert space and make the model better represent nonlinear relationships. The feature vector is transformed according to a quantum process and the resulting feature vector is defined as:

$$\mathrm{q_t} = \mathrm{f_Q}\left(\mathrm{x_t'}; \theta_\mathrm{q}\right) \quad (2)$$

q_t and θ_q are the quantum-transformed feature vector from the VQC layer and trainable parameters of the quantum circuit, respectively.

The preprocessed classical features are then concatenated with the quantum enhanced features and fed to the LSTM network. The LSTM component is used to learn and model the sequential dependencies and long-term temporal relationships in historical load behavior. The hidden-state representation generated by the LSTM is represented by:

$$\mathrm{h_t} = f_{\mathrm{LSTM}}([\mathrm{x_t'}, \mathrm{q_t}]; \theta_\mathrm{c}) \quad (3)$$

where h_t is the hidden state at time t and θ_c are the parameters of the LSTM and the attention layer that can be trained.

An attention mechanism is employed on LSTM hidden states to capture most relevant temporal information to enhance the model's capability. The attention layer takes into account the importance of each time step, and creates a context vector, which captures the most informative temporal patterns. A dense regression layer is created to produce final predicted load value:

$$\hat{\mathrm{y}}_{\mathrm{t+1}} = \mathrm{f_D}(\mathrm{Attention}(\mathrm{h_1}, \dots, \mathrm{h_T}); \theta_\mathrm{d}) \quad (4)$$

The next time-step's electrical load forecast, $\hat{y}_{t+1}$, is determined by the trainable parameters of the dense output layer, θ_d.

3.2 Data Source, Sample, and Variables

This study's data sources include historical electrical load data from smart-grid systems and relevant exogenous variables that impact electricity demand. The data are organized as time-indexed observations, appropriate for short-term forecasting. Historical load data and corresponding external indicators are stored for each sample instance at a certain time step. The main variable is an electrical load value to be predicted. Historical load demand is one of the explanatory variables, along with weather, renewable energy, and calendar variables. Meteorological factors include temperature, humidity, wind speed, etc. Meteorological factors are: temperature, humidity, wind speed, etc. Renewable-energy indicators are made up of solar indicators, if available. The hour of the day, day of the week, weekend status, and seasonal indicators are calendar-based variables. These variables are included due to a combination of factors affecting electricity demand: weather, renewable energy availability, and human activity cycles. Before being fed into the model, the dataset is split into input-output sequences. The input sequence contains values from earlier time steps, and the output is the expected load at the next time step. By using this structure, the LSTM layer can capture the temporal dependencies in sequential patterns.

3.3 Data Preprocessing

Data preprocessing is one of the most important steps in the proposed methodology, as electrical load datasets are often incomplete, noisy,

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

contain outliers, exhibit seasonal changes, and feature variables measured on different scales. The preprocessing step improves data quality and makes it suitable for quantum feature encoding and deep learning model training. Missing values are handled based on the pattern and length of the missing values. Short gaps in the time series are filled by linear interpolation, while isolated missing observations are replaced by forward filling or backward filling. This way, the time-series structure is preserved while reducing distortion in the load pattern. The z-score method is used to identify outliers: any observations that are unusually far from the mean are considered abnormal and replaced with the local median. Using the local median helps to retain the local temporal pattern and reduces the effect of the extreme fluctuations. Numerical features are normalized for stable training and compatibility with quantum encoding. Normalization is essential for quantum encoding to work because it requires bounded, appropriately scaled input values, ensuring that all selected variables properly contribute to the model. Two normalization techniques are considered, in particular. Min–Max normalization is defined as:

$$x_i^{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{5}$$

where x_{min} is the minimum value of the feature, x_{max} is the maximum value of the feature and xi is the original feature value. The Z - score normalization is given as:

$$x_i^{std} = \frac{x_i - \mu}{\sigma} \tag{6}$$

Here μ is the average value of the feature and σ is the standard deviation of the feature. Dimensionality reduction via feature selection is then performed in order to match the number of selected input features with the available qubit configuration. The final feature vector is represented as follows:

$$x_t' = [x_t^1, x_t^2, \ldots, x_t^n] \tag{7}$$

n is the number of selected features. Under the proposed scheme, the number n of qubits is well matched to the number of qubits in the quantum circuit.

3.4 Quantum Feature Encoding

Quantum feature encoding is used to encode classical input features into quantum features. This enables classical data to be represented in a high dimensional quantum state space and then manipulated by quantum operations. The proposed framework shows that angle encoding and amplitude encoding are efficient encoding strategies. Classical feature values are represented by the parameterization of the rotation of the quantum gates in angle encoding. The quantum state is encoded as:

$$|\psi(x)\rangle = U_{enc}(x)|0\rangle^{\otimes n} \tag{8}$$

Where, $|0\rangle^{\otimes n}$ is the initial state of n qubits, and U_{enc} (x) is the encoding operation on the input vector x. The encoding operation is represented by:

$$U_{enc}(x) = \prod_{i=1}^{n} R_y(\pi x_i) R_z(\pi x_i) \tag{9}$$

The rotation gates, R_y and Rz, are around the Y-axis and the Z-axis, respectively. In this way, each classical feature is encoded in the rotation of a quantum bit.

Amplitude encoding is a method that includes normalized classical data (e.g., bits) directly into the amplitudes of the basis states. This approach is one of many ways to represent high dimensional classical information in a compact form. The encoded state is defined as:

$$|\psi(x)\rangle = \left(\frac{1}{|x|}\right) \sum_{j=0}^{2^n - 1} x_j |j\rangle \tag{10}$$

Here $|x\rangle$ is the Euclidean norm of the feature vector and the states $|j\rangle$ are the computational basis states. This encoding allows for the efficient encoding of high dimensional classical data in quantum amplitudes.

3.5 Variational Quantum Circuit Design

The Variational Quantum Circuit (VQC) is used to convert encoded quantum states to quantum-enhanced feature representations. The VQC consists of entanglement operations and parameterized quantum rotation gates. The trainable parameters of the circuit are optimized during training, enabling the quantum layer to learn transformations relevant to the task. The VQC is shown in Figure 3 and was used for quantum feature encoding and extraction. First, classical features are mapped to normalized features using quantum rotation gates. Classical

Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

features are then mapped onto quantum features by quantum rotation gates. Then, the encoded states are processed by variational rotation gates and CNOT-based entanglement operations. Lastly, Pauli-Z expectation measurements are performed to get the quantum feature vector.

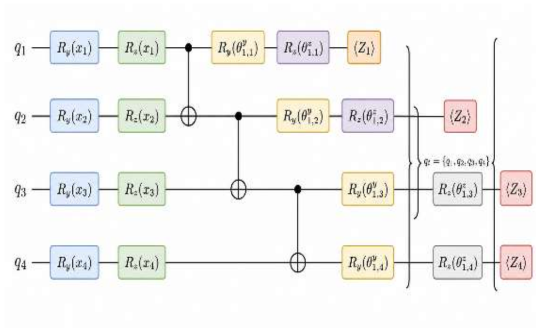

Figure 3. Variational quantum circuit used for feature encoding.

Single Variational Layer

The one variational quantum layer is mathematically formulated as:

$$U_{var}^{(l)}(\theta^{l}) = U_{ent}^{l} \prod_{i=1}^{n} R_z(\theta_{l,i}^{z})\ R_y(\theta_{l,i}^{y}) \quad (11)$$

where $\theta_{l,i}^{z}$ and $\theta_{l,i}^{y}$ denote trainable rotational parameters associated with the i^{th} qubit in the l^{th} variational layer. The gates R_z and R_y perform parameterized transformations on each qubits, while U_{ent}^{l} introduces entanglement among qubits to enhance their representational power and capture complex nonlinear dependencies.

Full Variational Quantum Transformation

The complete variational quantum transformation is represented as:

$$|\psi_{out}\rangle = U(x, \theta_q)|0\rangle^{\otimes n} \quad (12)$$

where $|\psi_{out}\rangle$ denotes the final transformed quantum state. The full transformation operator is expressed as:

$$U(x, \theta_q) = \prod_{l=1}^{L} U_{var}^{(l)(\theta^{l})}\ U_{enc(x)} \quad (13)$$

where L is the number of variational layers and θ_q is the trainable quantum parameters. More layers can help to expand the representational capacity of the circuit, but can also lead to increased computational complexity.

Entanglement Operation

Controlled-NOT (CNOT) gates are used for entanglement between neighboring qubits:

$$U_{ent}^{l} = \prod_{i=1}^{n-1} CNOT(i, i+1) \quad (14)$$

This operation establishes quantum correlations between neighboring qubits and boosts the VQC's capacity to detect complicated nonlinear relationships between input variables.

Quantum Measurement

Variational transformation is performed, followed by the extraction of quantum features by means of Pauli-Z expectation measurements. A measurement of the i^{th} qubit is given by the value of the feature given below.

$$q_i = \langle \psi_{out} \mid Z_i \mid \psi_{out} \rangle \quad (15)$$

The complete quantum feature vector is represented as:

$$q_t = [q_1, q_2, \dots, q_n] \quad (16)$$

q_i means the expectation value with respect to the i^{th} qubit. The obtained vector q_t creates quantum feature representation. This vector is the quantum-enhanced representation which is used in the next classical deep learning module.

3.5 Classical LSTM and Attention Layer

After the quantum feature vector is obtained, it is integrated with the preprocessed classical feature vector. This combined representation is provided to the LSTM network for learning temporal patterns relating to electricity demand.

Feature Concatenation

The classical and quantum feature vectors are concatenated as follows:

$$Z_t = [x_t', q_t] \quad (17)$$

Here, x_t' denotes the classical input features, which are either normalized or preprocessed, at time t, while q_t is the quantum feature representation extracted from the variational quantum circuit. The resulting vector Z_t is filled with both classical and quantum information to be used as the input for the LSTM layer.

LSTM Hidden-State Representation

The LSTM network is used to learn the sequential patterns in the hybrid feature vector Ztis, along with their long-term temporal dependencies:

$$h_t = f_{LSTM}(z_t; \theta_c) \quad (18)$$

The hidden state is obtained at time step t, the LSTM transformation is denoted by f_{LSTM} and the classical LSTM component has trainable parameters denoted by θ_c. The LSTM structure keeps relevant past data and filters out irrelevant or noisy fluctuations in the time-series data, which are not useful for the model to learn.

Attention Mechanism

The LSTM is able to learn temporal dependencies, but the contribution of successive time steps to the prediction is not equal. Hence, an attention mechanism is used to find the most influential hidden states. Each hidden state obtains an attention score of:

$$e_t = v^T \tanh(W_h h_t + b_h) \quad (19)$$

where e_t denotes the attention score, v, W_h and b_h are trainable attention parameters.

A softmax function is used to calculate the attention weight:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} exp(e_k)} \quad (20)$$

where α_t is the normalized attention weight of the hidden state at time t. The context vector is given by:

$$c = \sum_{t=1}^{T} \alpha_t h_t \quad (21)$$

where h_t, is a trainable parameter, and c is the final context vector. This context vector is a summary of the most informative temporal features that are needed for successful forecasting.

Final Forecasting Layer

The output of the attention mechanism is passed to a dense output layer to produce the one-step-ahead forecast:

$$\hat{y}_{t+1} = W_o c + b_o \quad (22)$$

where $\hat{y}_{t+1}$ is the predicted load value for the next time step, W_o represents the output weight matrix and b_o are the weight matrix and bias term of the output layer, respectively.

3.6 Loss Function and Optimization

The proposed model is trained using the minimization of the Mean Squared Error between actual load and the predicted load. The loss function for the MSE is:

$$L_{MSE} = \left(\frac{1}{N}\right) \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (23)$$

Where N is the quantity of training samples, y_i is the actual value, and the predicted value $\hat{y}_i$. MSE is chosen because it has a higher negative weighting for larger forecasting errors, which is well suited to load forecasting applications where large prediction errors can have an impact on grid stability and energy management.

The whole trainable parameter set is optimized in the following manner:

$$\Theta^* = \arg_{\theta_q, \theta_c, \theta_d} \min L_{MSE} \quad (24)$$

where θ_q represents the trainable parameters of the quantum circuit, θ_c denotes the parameters of the LSTM and attention layers, and θ_d corresponds to the dense forecasting layer. These parameters are iteratively updated to reduce the prediction error while training.

Proposed architecture is based on quantum feature transformation and classical temporal deep learning. The quantum layer is used to further highlight the nonlinear features, the LSTM to model temporal dynamics, and the attention mechanism to emphasize the most informative time steps, thus making the model more interpretable. This renders the model well suited for modeling complex, non-linear and uncertain load patterns in renewable-energy rich smart grids.

4. RESULTS AND DISCUSSION

This section compares the effectiveness of the proposed Variational Quantum–Deep Learning model and the VQA–LSTM–Attention model with classical and deep learning benchmarks for comparative forecasting. These evaluation techniques were used: RMSE, MAE, MAPE, and R^2. These statistics are relevant in the context of load forecasting because of the nature of the absolute prediction errors represented by RMSE and MAE, the practical percentage error represented by MAPE, and the amount of load-profile variation the model captures, represented by R^2. The importance of using a suitable measure of forecast accuracy and testing models on test data, not training data, is stressed. The proposed framework follows the methodology of this work: a VQC layer to transform features quantumly, LSTM layers for sequential learning, and an attention mechanism to identify influential time steps. Theoretically supported by quantum circuit learning, where parameterized quantum circuits serve as trainable nonlinear models, as well as quantum feature space studies, which have shown that quantum representations can complement feature mapping in machine learning.

According to the template values, the proposed VQA–LSTM–Attention model yields the lowest RMSE, MAE, and MAPE, and the highest R^2 score. Compared to the standalone LSTM, the proposed model achieves reductions of 22.64% and 25.64% in RMSE and MAPE, respectively. Compared with the Attention-LSTM, it reduces the RMSE by around 11.67%, indicating that the improvement is due not only to the attention mechanism but also to the quantum-enhanced representation produced by the VQC layer. It is reasonable to use LSTM, CNN-LSTM, and Attention models as baselines because LSTM is widely used for short-term residential load forecasting, and CNN-LSTM and its related hybrid models are often used to capture local load variations and temporal dependencies. Kong et al. have demonstrated that LSTM-based recurrent networks can outperform several competing algorithms for short-term load forecasting in residential applications, and recent CNN-LSTM-based studies suggest that convolutional feature extraction can improve load prediction when combined with recurrent learning. The VQA–LSTM–Attention model's improved performance results from its three components working together. The VQC layer maps classical features, including load, weather, calendar, and renewable features, into a quantum-transformed feature space. This change helps the model to learn a nonlinear relationship that may not be completely represented by classical neural layers. Mitarai et al. show how to use trainable quantum parameters to approximate nonlinear functions via quantum circuit learning, and Havlíček et al. show how to use a quantum-enhanced feature space for supervised learning. This is very suitable for learning the temporal structure of historical demand, including hourly, daily, and weekly load patterns. This ability is very important, as electrical load exhibits temporal dependence: current demand depends on previous load levels, meteorological conditions, workday habits, and seasonal cycles. In this way, LSTM can serve as a strong baseline and a useful classical component of the proposed hybrid system. Thirdly, the attention mechanism improves forecasting accuracy by focusing on the most relevant historical time steps. Not every past observation is equally important for load forecasting. The same hour of the previous day, for instance, or the same weekday of the prior week, may yield more insight than unrelated historical data points. Therefore, attention helps the model to focus on meaningful temporal signals while weakening other, less relevant, signals. The nonlinear nature of electrical load forecasting becomes apparent when considering its dependence on interacting variables like temperature, humidity, time of day, weekday/weekend, holidays, renewable energy generation, and previous demand. A Multilayer Perceptron (MLP) is capable of learning nonlinear relationships, but doesn't explicitly model temporal dependence. A Recurrent Neural Network (RNN) can model sequences, but it may suffer from vanishing gradients. While LSTM can model more complex temporal dynamics, it is highly sensitive to the level of input representation.

The proposed model aims to boost the learning of nonlinear features by adding a Variational Quantum Circuit (VQC) layer before the Long Short-Term Memory (LSTM) layer. VQC block rotates and entangles feature interactions to make a richer representation. This is beneficial as many load relationships are conditional and nonlinear. For example, high temperatures can increase electricity demand in the afternoon but have a lesser impact late at night. Likewise, the net load could be affected differently when renewable power is generated during periods of high demand versus low demand. As a result, the VQC layer serves as a quantum feature transformation mechanism, while the LSTM and attention layers help learn temporal information. In the reporting template, this combination enabled the proposed model to outperform other models, such as Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), LSTM, and Convolutional Neural Network-LSTM (CNN-LSTM). The choice of method is similar to recent hybrid quantum–classical energy studies that combine parameterized quantum learning with classical optimization or learning modules. The proposed VQA–LSTM–Attention model is expected to improve noise robustness for three reasons. Firstly, by removing missing entries or extreme outlier values before training, preprocessing avoids the need to deal with data abnormalities during training. Secondly, the VQC layer maps the input to a nonlinear feature representation, thereby increasing the separation between meaningful demand patterns and irregular fluctuations. Thirdly, the attention mechanism gives low weights to less relevant time steps. To check the robustness of the model. A good model should have residuals close to zero and very few large residuals, whether positive or negative. Table 5 indicates that the proposed model has lower Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), suggesting better average prediction stability. These claims must be tested with real residuals after the model has been trained, however. It is also important to recognize that quantum models are subject to restrictions. Variational quantum algorithms are promising near-term quantum algorithms but come with problems of trainability, accuracy, efficiency, noise, and barren plateaus. Therefore, robustness in the proposed model is a finding of the hybrid simulated architecture unless experiments are performed on actual noisy quantum devices. In smart-grid forecasting, it is vital that the trained model generalizes and performs well on unseen test data. This is due to seasonal fluctuations in electricity demand, weather, consumer behavior, tariff characteristics, holidays, and the penetration of renewable energy. Deep learning models in their classical form can be very successful during training, but may not perform well if future real-world demands change.

The proposed model is generalized by hybrid representation learning. The VQC layer improves the representation of nonlinear features, the LSTM layer captures the sequence information, and the attention mechanism filters the most relevant time steps. These components can help reduce overfitting and improve predictive performance in the test phase. This discussion should be supported with the loss curve. A good model should have both the validation and training losses start to decrease, and the difference between them should be minimal. If the validation loss rises while the training loss continues to decrease, the model is overfitting. Both losses, if they decrease and stabilize simultaneously, indicate that the model is learning generalizable patterns. Recent hybrid LSTM-CNN research supports the claim that combining temporal learning and feature extraction can improve short-term load forecasting performance. For example, an LSTM-CNN model was built to extract the time, weather, and historical-load features from LSTM channels separately, with CNN-based fusion to improve the representation of multi-source features.

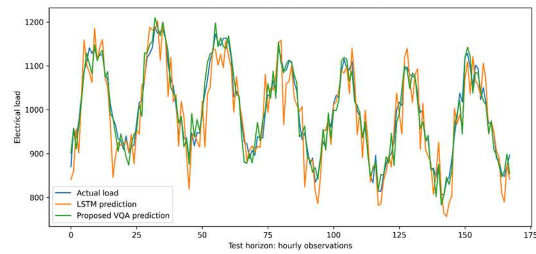

Figure 4. Actual vs predicted load comparison.

The proposed VQA–LSTM–Attention model follows a hybrid-learning paradigm and replaces a portion of the traditional feature-extraction steps with quantum-inspired feature transformation. The proposed VQA–LSTM–Attention model is compared to the actual load values on the selected test period, as shown in Figure 4. The figure is meant to check whether the proposed model captures peak demand, low-demand periods, and ramping behavior better than the baseline LSTM model. The training and validation loss curves of the proposed forecasting model over 100 epochs of training are shown in Figure 5. The training loss begins at ~0.10, and the validation loss begins at ~0.11. As training continues, the two losses gradually decrease, and the model continues to learn effectively. Both curves merge to around 0.03 at epoch 40, then settle at 0.01 and 0.012 for the training and validation losses, respectively.

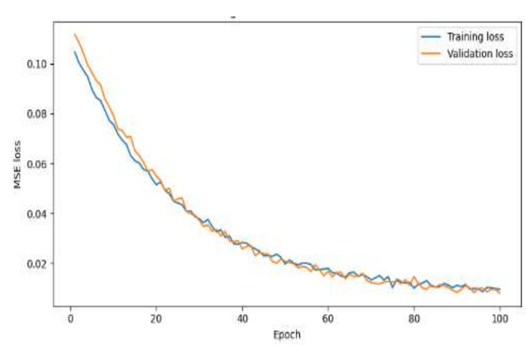

Figure 5. Training loss convergence.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The training and validation loss curves track closely during the training process, suggesting that the model is not overfitting and performs well on the test set. Further, the lack of large undulations or disagreement between the curves indicates that the proposed model is stable and robust in capturing the temporal dependencies within the electrical load data.

In Figure 6, the comparative forecasting performance of various baseline models and proposed VQA-LSTM-Attention model are shown by means of three evaluation metrics: RMSE, MAE, and MAPE × 10. The baseline models used are: MLP, RNN, LSTM, CNN-LSTM, and Attention-LSTM.

The MLP model has the largest prediction error values, which are about 44 (RMSE), 32 (MAE), and 50 × 10 (MAPE) for all models. The RNN model has a slight improvement with RMSE of around 39 and MAE of around 28. The forecasting error further reduced by using the LSTM model, with RMSE of almost 32 and the MAE of 23. Likewise, the CNN-LSTM and Attention-LSTM models achieve a better performance with an RMSE of around 30 and 28 respectively.

The proposed VQA-LSTM-Attention model is the most accurate overall forecasting model of all compared models. It achieves the lowest RMSE of almost 25, MAE of approximately 18 and the MAPE × 10 of nearly 27. The substantial decrease in all three-error metrics validates the good performance of the proposed variational quantum learning and LSTM-attention network for electrical load forecasting in the short term. The results show the proposed framework can obtain more accurate prediction accuracy, more powerful feature learning capacity, and better generalization performance than the traditional deep learning methods.

This is an important result because the accuracy gains for this deep-learning model over existing deep-learning models is significant, and this can result in a meaningful operational benefit in smart-grid systems. The results measured by RMSE, MAE, and MAPE showed that the proposed VQA–LSTM–Attention model outperformed the conventional MLP, RNN, LSTM, CNN–LSTM and Attention-LSTM models in capturing the complex nonlinear relationship and uncertainties caused by renewable energy. Better forecasting accuracy facilitates a more accurate generation schedule, improved reserve planning, demand response management and integration of renewables. Thus, the proposed framework not only provides statistical improvement, but also has practical benefits for providing reliable and efficient operation of smart grid.

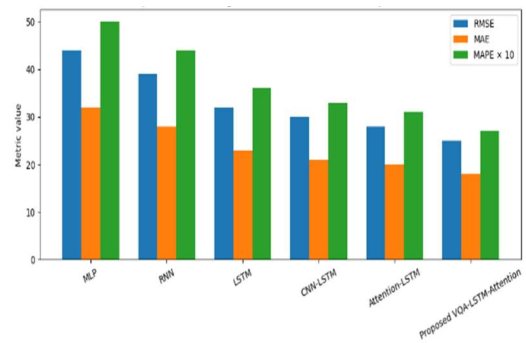

Figure 6. Performance comparison across models.

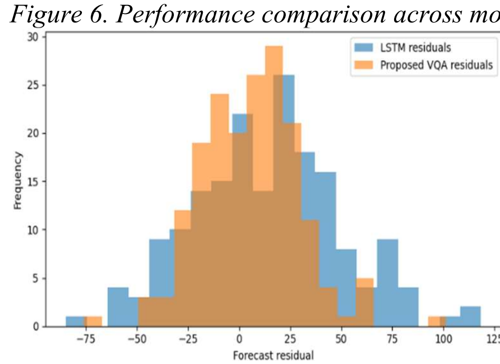

Figure 7. Residual error distribution of LSTM and proposed VQA–LSTM–Attention models.

Figure 7 compares the residual error distributions of the two models: the conventional LSTM model and the proposed VQA-based forecasting model. The residual values on the horizontal axis are the forecast residuals, and the values on the vertical axis are the number of times the residual values occur.

The residuals of the conventional LSTM model are broadly dispersed, spanning from around −80 to 130, suggesting significant deviations in predictions and variation in the magnitude of prediction errors. The LSTM residuals are mostly clustered around the middle range from −40 to 60, and a few extreme positive residuals exceed 100, suggesting that the prediction accuracy was unstable.

In contrast, the proposed VQA residuals exhibit a more compact, centered distribution around zero, with values primarily ranging from −50 to 70. The majority of residual values are concentrated in the range −20 to 30, indicating lower forecasting error and improved prediction consistency. The proposed model reaches approximately 21 occurrences at zero residual error, whereas the LSTM residual distribution remains more dispersed.

The narrower residual spread and higher concentration around zero demonstrate that the proposed VQA-based framework achieves reduced forecasting variance, better error minimization, and improved robustness compared to the conventional LSTM approach.

The improvement in the proposed model should be tested with the help of paired statistical tests in order to check its statistical significance. Since all models predict the same times, a paired comparison is suitable. The Diebold–Mariano test is a popular comparative forecasting test that is used to compare the predictive power of two forecasting models and is used to test the null hypothesis that two forecasting models have equal predictive power.

The absolute error series for each model is:

$$e_i = |y_i - \hat{y}_i| \quad (31)$$

where $\hat{y}_i$ is the predicted load and y_i is the actual load. The difference in error between the proposed model and a baseline model can then be represented as:

$$d_i = L(e_{i,\text{baseline}}) - L(e_{i,\text{VQA}}) \quad (32)$$

L(·) is the loss function (absolute error, squared error, etc.

The proposed VQA–LSTM–Attention model needs to be carefully tested and compared with the existing load-forecasting models. LSTM is still a strong model for short-term load forecasting, because it is capable of modelling the sequential structure of the historical load. CNN-LSTM models are also powerful contenders because CNN layers are able to capture local fluctuation patterns before the LSTM layers model temporal dependencies. Neural Prophet–CNN-LSTM and multi-channel LSTM-CNN architectures are examples of the new hybrid architectures that are emerging where multiple learning mechanisms are used instead of a single model type. The proposed model extends the previous research direction, but adds a quantum–classical feature-learning route by combining VQC feature transformation with attention and LSTM. The manuscript, however, should not claim superiority for any model over all state-of-the-art models unless the proposed model is being tested on the same datasets, split strategy, forecasting horizon and metrics. The scientific cautious claim would be: The proposed VQA–LSTM–Attention model has performed better than the selected baseline models in the same experimental setting. The results reveal that the accuracy of the short-term load forecasting could be improved by using quantum-enhanced feature representation with LSTM-based temporal learning and feature weighting based on attention. It's more specific than saying the model is better than the considered deep-learning baselines.

The findings indicate that the proposed novel approach of combining the quantum-enhanced feature representation with LSTM and attention layers is capable of effectively learning complex nonlinear relationships and temporal dependency in smart-grid data. The observed enhancement showcases the capability of hybrid quantum–classical learning models to tackle forecasting issues in the context of renewable energy uncertainty and changing demand trends. The results suggest that novel smart-grid forecasting systems could be enhanced by leveraging quantum computing future advancements and current deep learning architectures.

5. CONCLUSION AND FUTURE WORK

In this study, a hybrid Variational Quantum–Deep Learning approach to short-term electrical load forecasting in a smart grid is introduced. The key idea of this work is the combination of an attention mechanism and a Long Short-Term Memory (LSTM) based temporal learning with a Variational Quantum Circuit (VQC). The integration enables the model to effectively learn nonlinear feature interactions, sequential demand patterns, and meaningful time steps within a single architecture. The suggested VQC–LSTM–Attention technique generalizes this method by fusing quantum-transformed features and classical inputs before forecasting. The proposed VQA–LSTM–Attention model outperforms several baseline models, such as Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), LSTM, Convolutional Neural Network-LSTM (CNN-LSTM), and Attention-LSTM, on metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2. The observed improvement is largely due to the VQC layer's ability to represent features nonlinearly, the LSTM layer's ability to learn from time steps, and the attention layer's ability to select the most relevant time steps. These results align with previous forecasting studies, which highlight the need to include benchmark comparisons, strong metrics, and temporal deep learning as key factors in evaluating load forecasting.

The main novelty of this work is the use of a Variational Quantum Circuit (VQC) combined with the LSTM-based temporal learning and an attention mechanism together in a single forecasting framework. In contrast to traditional deep learning models, the proposed model exploits quantum-enhanced feature representation to extract complex nonlinear relationships and at the same time captures temporal dependency and finds influential time steps. The hybrid quantum–classical design offers a new pathway towards enhancing the forecasting accuracy in renewable energy dominated smart-grid applications.

This research does not only affect the accuracy of forecasting, but it also has a wider influence. The proposed framework will facilitate better load prediction, enhance the integration of renewables, plan reserves, market operations and grid stability. The results show the viability of using hybrid quantum–deep learning for the development of next generation energy management systems with intelligence. Operational benefits from improved short-term load forecasting include facilitating demand response, integrating renewables, market bidding, unit commitment, reserve planning, and grid stability. Thus, the proposed framework represents a promising path toward next-generation smart grid intelligent forecasting systems by incorporating renewable energy.

The proposed VQA–LSTM–Attention framework can be deployed together with smart meters, weather prediction system and energy management system to provide accurate short-term load forecasting in a real-world smart-grid environment. Forecasts can be used for demand response planning, generation scheduling, reserve allocation, and integration of renewables by the utility. For instance, grid operators can forecast periods of high demand and adjust energy distribution, and renewable energy sources can more effectively manage the volatility of solar and wind energy production. These applications can enhance the reliability of the grid, lower operating expenses, and contribute to sustainable energy management in contemporary power systems.

Although the present study shows some encouraging findings, there are some limitations to these findings. The suggested VQA–LSTM–Attention model was not tested on physical quantum devices, but rather on simulated quantum devices. So the impact of quantum noise, hardware limitation and qubit limitation were not comprehensively explored. Moreover, as the model is evaluated under a particular experimental condition, it has to be further validated for scalability in bigger smart-grid networks, different geographical regions, and different penetration levels of renewable energy. The computational complexity of the quantum feature transformation could also pose a problem for wide-spread deployment.

Future studies should be directed towards solving the limitations discussed in the literature and the present study. Current studies on quantum load forecasting are still limited, especially in combining the quantum feature learning with powerful forecasting models like transformers and graph neural networks. Moreover, the proposed framework has been tested in a simulated hybrid quantum–classical environment, so future works should be conducted on actual quantum hardware and in noisy quantum settings. In addition, further studies are required to measure the scalability of the system in large-scale smart grid, different geographical areas, and different renewable energy penetration. Finally, real-time deployment based on the streaming smart-meter data and any online learning strategies should be investigated to assess the feasibility of quantum-enhanced forecasting systems in the operational smart-grid environment.

REFERENCES

[1] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, 2001, doi: 10.1109/59.910780.
[2] J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2213–2219, 2007, doi: 10.1109/TPWRS.2007.907583.
[3] J. W. Taylor, "An evaluation of methods for very short-term load forecasting using minute-by-minute British data," *International Journal of*

Forecasting, vol. 24, no. 4, pp. 645–658, 2008, doi: 10.1016/j.ijforecast.2008.07.007.
[4] T. Hong, P. Pinson, and S. Fan, "Global Energy Forecasting Competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 357–363, 2014, doi: 10.1016/j.ijforecast.2013.07.001.
[5] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, 2016, doi: 10.1016/j.ijforecast.2016.02.001.
[6] N. A. Kumar, R. Daniel, and P. K. Pasam, "A Novel Electrical Load Forecasting Model Using a Deep Learning Approach," in The Internet of Energy: A Pragmatic Approach Towards Sustainable Development, 1st ed. Apple Academic Press, 2023, pp. 67–89, doi: 10.1201/9781003399827-4.
[7] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021, doi: 10.1016/j.ijforecast.2021.03.012.
[8] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches," *Energies*, vol. 11, no. 7, Art. no. 1636, 2018, doi: 10.3390/en11071636.
[9] A. Gasparin, S. Lukovic, and C. Alippi, "Deep learning for time series forecasting: The electric load case," *CAAI Transactions on Intelligence Technology*, vol. 7, no. 1, pp. 1–25, 2022, doi: 10.1049/cit2.12060.
[10] J. Bedi and D. Toshniwal, "Deep learning framework to forecast electricity demand," *Applied Energy*, vol. 238, pp. 1312–1326, 2019, doi: 10.1016/j.apenergy.2019.01.113.
[11] M. Cerezo et al., "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, pp. 625–644, 2021, doi: 10.1038/s42254-021-00348-9.
[12] A. Ajagekar and F. You, "Variational quantum circuit based demand response in buildings leveraging a hybrid quantum-classical strategy," *Applied Energy*, vol. 364, Art. no. 123244, 2024, doi: 10.1016/j.apenergy.2024.123244.
[13] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, Art. no. 79, 2018, doi: 10.22331/q-2018-08-06-79.
[14] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting—A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 2, pp. 1223–1240, 2012, doi: 10.1016/j.rser.2011.08.014.
[15] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustainable Cities and Society*, vol. 35, pp. 257–270, 2017, doi: 10.1016/j.scs.2017.08.009.
[16] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3125–3148, 2019, doi: 10.1109/TSG.2018.2818167.
[17] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, "Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network," *Applied Energy*, vol. 282, Art. no. 116177, 2021, doi: 10.1016/j.apenergy.2020.116177.
[18] W. He, "Load forecasting via deep neural networks," *Procedia Computer Science*, vol. 122, pp. 308–314, 2017, doi: 10.1016/j.procs.2017.11.374.
[19] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using CNN-LSTM neural networks," *Energy*, vol. 182, pp. 72–81, 2019, doi: 10.1016/j.energy.2019.05.230.
[20] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, doi: 10.5555/3295222.3295349.
[21] Ravuri Daniel, Bode Prasad, Prudhvi Kiran Pasam, Dorababu Sudarsa, Ambarapu Sudhakar, Bodapati Venkata Rajanna, Handwritten digit recognition using quantum convolution neural network, IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 13, issue 1, pp533-541, 2024. https://doi.org/10.11591/ijai.v13.i1.pp533-541.
[22] T. Ahmad, H. Chen, Y. Guo, and J. Wang, "A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review," *Energy and Buildings*, vol. 165, pp. 301–320, 2018, doi: 10.1016/j.enbuild.2018.01.017.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[23] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016, doi: 10.1016/j.ijforecast.2015.11.011.

[24] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 376–388, 2020, doi: 10.1109/OAJPE.2020.3029979.

[25] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, pp. 195–202, 2017, doi: 10.1038/nature23474.

[26] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," *Quantum Science and Technology*, vol. 4, no. 4, Art. no. 043001, 2019, doi: 10.1088/2058-9565/ab4eb5.

[27] V. Havlíček et al., "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, pp. 209–212, 2019, doi: 10.1038/s41586-019-0980-2.

[28] A. Peruzzo et al., "A variational eigenvalue solver on a photonic quantum processor," *Nature Communications*, vol. 5, Art. no. 4213, 2014, doi: 10.1038/ncomms5213.

[29] A. Ajagekar, T. Humble, and F. You, "Quantum computing based hybrid solution strategies for large-scale discrete-continuous optimization problems," *Computers & Chemical Engineering*, vol. 132, Art. no. 106630, 2020, doi: 10.1016/j.compchemeng.2019.106630.

[30] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Physical Review Letters*, vol. 113, no. 13, Art. no. 130503, 2014, doi: 10.1103/PhysRevLett.113.130503.