Journal of Renewable Energy and Sustainable Development (RESD)
Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

Received: 24 January 2026 Accepted: 19 April 2026 Published: 03 May 2026

# Quantum-based Hybrid Framework for Adaptive Congestion Management in Smart Grids with Stochastic Renewable Penetration

**Milan Sasmal [1*], Partha Das [2], Biswamoy Pal [3], Suman Ghosh [4], and Souvik Dutta [5]**

[1*,2,3] JIS College of Engineering, Electrical Engineering Department, Kalyani, West Bengal, India

[4] Guru Nanak Institute of Technology, Electrical Engineering Department, West Bengal, India

[5] Haldia Institute of Technology, Electrical Engineering Department, West Bengal, India

milan.sasmol@jiscollege.ac.in, partha.das@jiscollege.ac.in, biswamoy.paul@jiscollege.ac.in, suman.ghosh@gnit.ac.in, souvikdutta8808@gmail.com

## Abstract

*Integrating renewable energy sources (RES) into conventional power systems introduces significant stochastic volatility. This volatility often leads to frequent transmission line congestion and voltage instability in modern smart grids. Existing hybrid optimization strategies, typically combining classical Reinforcement Learning (RL) and Particle Swarm Optimization (PSO), are utilized for congestion management. However, these methods often encounter performance degradation due to the dimensionality of large-scale networks, such as the IEEE 118-bus system.*

*This paper presents a quantum-hybrid framework integrating Variational Quantum Reinforcement Learning with Quantum-behaved PSO (VQRL-QPSO). The architecture utilizes Variational Quantum Circuits (VQC) to map high-dimensional grid states into a Hilbert space via Angle Encoding. This approach enables the RL agent to learn optimal control policies using significantly fewer parameters than classical deep neural networks. Additionally, the QPSO component incorporates a wave-function-based search mechanism using a Delta potential well model. This mechanism facilitates "quantum tunneling," allowing the optimizer to escape local optima that typically entrap classical heuristics.*

*The IEEE 39-bus and 118-bus systems were utilized for simulation and validation. Results indicate that the VQRL-QPSO framework achieves an approximate 25% reduction in operational congestion costs by accelerating convergence time 48% compared to conventional RL-PSO benchmarks. Furthermore, grid resilience and voltage stability are maintained even under 50% renewable uncertainty. The proposed framework establishes the potential of quantum-hybrid algorithms as scalable, real-time tools for ensuring stability in high-penetration smart grids.*

**Index-words:** *Congestion Management, IEEE 118, QPSO, Quantum, Reinforcement Learning, Renewable Energy, Smart Grid, Variational Quantum Circuits.*

## I. Nomenclature

| Abbreviations | |
|---|---|
| **Term** | **Meaning** |
| CNOT | Controlled-NOT Gate |
| DQN | Deep Q-Network |
| IEEE | Institute of Electrical and Electronics Engineers |
| LMP | Locational Marginal Pricing |
| MSE | Mean Squared Error |
| NISQ | Noisy Intermediate-Scale Quantum |
| OPF | Optimal Power Flow |
| PCA | Principal Component Analysis |
| PDF | Probability Density Function |
| PSO | Particle Swarm Optimization |
| p.u. | Per-unit |
| PV | Photovoltaic |
| QPSO | Quantum-behaved Particle Swarm Optimization |
| VQC | Variational Quantum Circuit |
| VQRL | Variational Quantum Reinforcement Learning |
| RES | Renewable Energy Sources |
| RL | Reinforcement Learning |

* Corresponding author

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

| Power System Symbols | |
|---|---|
| **Term** | **Meaning** |
| C | Scale factor for the Weibull distribution |
| $C_{ong}$ | Total congestion cost derived from rescheduling |
| $G$, $G_{max}$ | Instantaneous and maximum solar irradiance |
| $G_{ij}$, $B_{ij}$ | Conductance and susceptance of the line between bus i and j |
| $J$ | Cost function |
| $L_k$ | Loading percentage of line k |
| k | Shape factor for Weibull distribution |
| $N_{itr}$ | Number of iterations |
| $T_s$ | Simulation time |
| $T_f$ | Power flow analysis time |
| $P_{D,i}$ , $Q_{D,i}$ | Real and reactive power demand at bus i |
| $P_i$, $Q_i$ | Real, and reactive power generation at bus i |
| $P_{i\,max}$ , $P_{i\,min}$ | Ramp-rate constrained generation limits |
| $P_{wind}(v)$ | Output power of wind turbine at wind speed v |
| $P_{PV}(G)$ | Active power output of the solar PV plant |
| $S_k$ | Apparent power flow on transmission line k |
| $S_{k\_max}$ | Thermal limit of transmission line k |
| $S_{raw}$ | Raw input vector (118 voltages + 186 line flows) |
| $V_i$, $\theta_i$ | Voltage magnitude and phase angle at bus i |
| $V_{ref}$ | Reference voltage magnitude (1.0 p.u.) |

| Quantum Reinforcement Learning Symbols | |
|---|---|
| **Term** | **Meaning** |
| $\|\psi(s)\rangle$ | Quantum state representing the encoded grid state s |
| $\langle\hat{Z}_i\rangle$ | Expectation value of the Pauli-Z operator for qubit i |
| $\pi_\theta(a\|s)$ | RL policy represented by the quantum circuit |
| $\lambda_{1,2,3}$ | Penalty weights for congestion, voltage, and losses |
| γ | Discount factor for future rewards |
| $\omega_{1,2,3}$ | Multi-objective function weights |
| μ | Mean vector of the training dataset |
| σ | Standard deviation vector of the training dataset |
| $C_t$ | Computational cost |
| $R_t$ | Scalarized reward function at time t |
| $U_{enc}$ | Unitary operator for Angle Encoding (Pauli-Z/Y rotations) |
| $U_{var}(\theta)$ | Variational quantum circuit with trainable parameters θ |
| $s$, $a$, $r$ | State, Action, and Reward in the Markov Decision Process |

| QPSO Optimization Symbols | |
|---|---|
| **Term** | **Meaning** |
| $\beta$ | Contraction-expansion coefficient tuned by VQRL |
| $u$ | Uniform random variable u ~ U (0, 1) |
| $\Delta P_G$ | Generator power rescheduling vector (action output) |
| $g_{best,t}$ | Global best position of the swarm at iteration t |
| H | Entropy |
| $m_{best}$ | Mean best position of the particle swarm |
| $p_{i,t}$ | Personal best position of particle i at iteration t |
| $x_{i,t}$ | Position of particle i at iteration t |

## II. Introduction

The energy sector is undergoing a rapid transition toward decarbonization, characterized by a shift from centralized fossil-fuel power plants to decentralized renewable energy sources (RES). While essential, this transition introduces significant technical hurdles in power grid control. Wind and solar energy are inherently unpredictable, exerting substantial pressure on transmission infrastructures originally designed for deterministic generation patterns [1].

Transmission line congestion represents a critical challenge in modern power systems. Congestion occurs when power demand exceeds the physical, thermal, or stability limits of a transmission line [2]. Such occurrences result in increased operational costs and higher electricity prices. If unmanaged, congestion can trigger cascading failures and massive blackouts, similar to the 2003 Northeast blackout in North America and the 2012 India crisis [3]. The intermittent nature of RES generation causes rapid fluctuations, reaching grid-breaking points much faster than traditional sources.

Historically, congestion management relied on deterministic mathematical models, including generation rescheduling and standard Optimal Power Flow (OPF) algorithms. However, these methods often struggle with the complex, nonlinear dynamics of modern smart grids. Meta-heuristic tools, such as Particle Swarm Optimization (PSO) and Genetic Algorithms (GA), have been explored as alternatives [4, 5]. Despite their popularity, these heuristics frequently converge on sub-optimal solutions or require excessive computation time as network complexity increases.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

Reinforcement Learning (RL) has emerged as a promising approach for dynamic, real-time decision-making in power systems [6, 7]. Hybrid RL-PSO frameworks have demonstrated potential by utilizing RL agents to guide the optimization search [8]. Nevertheless, as systems transition to complex networks like the IEEE 118-bus system, classical deep learning models encounter significant computational scalability issues in high-dimensional state spaces.

Quantum Computing (QC) provides a viable solution to these scalability challenges [9, 10]. Variational Quantum Reinforcement Learning (VQRL) allows for the modeling of grid sensitivities using significantly fewer parameters than classical architectures [11]. By leveraging quantum superposition and entanglement, VQRL agents can map correlations across smart grids more effectively [12]. Furthermore, the integration of Quantum-behaved Particle Swarm Optimization (QPSO) introduces a search mechanism based on wave functions. This approach enables particles to "tunnel" through high-cost barriers to identify global optima [13, 14].

This paper proposes a new quantum-classical hybrid framework (VQRL-QPSO) specifically designed for congestion management in grids with high renewable penetration. The contributions of this research are highlighted as follows:

- Development of a VQRL agent that utilizes high-efficiency Principal Component Analysis (PCA) and angle encoding to map complex grid states.
- Implementation of an adaptive QPSO engine that dynamically tunes its own contraction expansion parameters based on the learned RL policy.
- Formal mathematical integration of quantum-behaved search mechanisms with reinforcement learning for multi-objective grid optimization.
- Validation using IEEE 39-bus and 118-bus systems, demonstrating superior performance over classical benchmarks in terms of operational Cost and convergence speed.

## III. Literature review

Nowadays, managing transmission congestion in smart grids has seen a huge transition. This transition is primarily caused by two major incidents: firstly, the rapidly increasing demand for electricity, and secondly, the worldwide integration of Renewable Energy Sources (RES), which is noisy in nature. This section describes how our proposed methodologies have evolved and their ability to stay steady despite the "diverse swings" of green energy.

In the previous days of power engineering, we treated the congestion as an Optimal Power Flow (OPF) challenge. The fundamental techniques established by Andersson [1], like the Newton-Raphson and interior point methods, fulfill the industry standard for keeping power flows within their lanes. As markets deregulated, Shahidehpour et al. [2] considered these ideas further by introducing Locational Marginal Pricing (LMP) as a financial way to keep the grid in check and stable. However, as Khator and Leung [15] pointed out, these models refer to a "perfect world" assumption where generation and load are predictable. That assumption simply is not acceptable in a world where RES like wind and solar can put a surge or drop in an instant.

Real-world transmission networks are more complicated, "nonlinear" and "nonconvex," so researchers eventually moved toward "meta-heuristic" optimization. Bio-inspired methods like Particle Swarm Optimization (PSO) by Kennedy and Eberhart [16] became popular due to the simplicity in terms of code and use. Similarly, Bijwe and Viswakarma [17] highlighted that Genetic Algorithms (GA) were quite robust at finding rescheduling plans. On the other hand, as highlighted by Patel et al. [4] and Wang et al. [5], these classical heuristics tend to get "trapped" in suboptimal solutions too early. When implemented in a large 118-bus system, due to the computational weight and the "noise" from renewable fluctuations, it often leads to a very slow performance.

The introduction of Artificial Intelligence was a major turning point for grid operations. "Reinforcement Learning" (RL), pioneered by Sutton and Barto [6], provides a path for the systems to take dynamic decisions by learning from their

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

environment, and later on, the same is refined for control systems by Busoniu et al. [7]. In the concept of smart grid, researchers like Li et al. [8] and Lee et al. [18] have also used Deep Q-Networks (DQN) to handle congestion in real-time. The hybrid RL-PSO approach was a natural next step, where the RL agent acts as the "strategist" and the PSO acts as the "searcher". But even this, Kumar et al. [19] noted that if more renewable energy is connected with the system, the state-space dimensionality for the RL agent grows exponentially, leading to the "curse of dimensionality" that slows down convergence in complex, multibus grids.

The unpredictable behavior of wind and solar requires advanced stochastic modeling. Researchers like Milano [20] and Brown et al. [21] have investigated the impacts of RES intermittency on grid stability. For a long time, operators relied on wide "safety margins", which underutilize renewable capacity. Now, in the age of the internet, we need adaptive and data-driven controllers that can "perceive" weather patterns in real-time to manage the grid more tightly.

Quantum Computing (QC) offers a potential path to scalability in comparison with classical AI. Schuld [12] proved that Variational Quantum Circuits (VQC) can approximate complex functions with a few parameters. Furthermore, Havlicek et al. [22] demonstrated that quantum-enhanced features can identify patterns in data that a classical computer would miss entirely. When we consider optimization, Sun et al. [14] introduced "Quantum-behaved" PSO (QPSO), which uses the wave-function principle to ensure global search capability. The application of QPSO to the transmission networks was further explored by Moradzadeh et al. [23], and it was established that "quantum tunneling" allows the optimizer to bypass the sub-optimal barriers created by strict thermal limits.

While the literature survey covers standard OPF, "meta-heuristic" PSO, and deep learning, there is a gap when it comes to merging Variational Quantum Reinforcement Learning (VQRL) with Quantum-behaved PSO for multi-objective congestion management. Most of the existing research is either focused on pure quantum optimization or pure classical learning. Our work put the steps into this gap by proposing a hybrid VQRL-QPSO framework. We aimed to combine the parameter efficiency of quantum circuits with the "tunneling" search capability of QPSO within the framework and to provide an optimal solution for congestion cost, grid voltage stability, and high penetration of RES.

## IV. System description & problem formulation

The IEEE 39-bus and IEEE 118-bus systems are utilized as the primary benchmarks for this study. The IEEE 118-bus system is a representative model of a large-scale power grid. It consists of 54 generators, 186 transmission lines, and 91 load centers. This network is highly nonconvex and possesses a vast search space for optimization [24-31].

High renewable penetration is simulated by integrating wind and solar farms at specific buses. Wind energy is injected at Bus 10, while solar photovoltaic (PV) generation is integrated at Bus 25. These stochastic sources introduce significant variability into the nodal power balance. The integration levels are varied up to 50% to test the robustness of the control framework.

Transmission line congestion occurs when the apparent power flow $(S_k)$ in a branch exceeds its rated thermal limit $(S_{k,max})$. In modern smart grids, this issue is exacerbated by the intermittent nature of renewable energy sources. Unpredictable surges in wind or solar output can lead to sudden violations of line capacity.

The objective of congestion management is to eliminate these violations through generator rescheduling. This process requires adjusting the active power output of conventional generators $(\Delta P_G)$ to reroute power flow. The primary challenges are ensuring line loading remains below 100%, maintaining bus voltage magnitudes within the standard [0.95, 1.05] p.u. range and minimizing the total Cost of rescheduling while satisfying demand.

Failure to manage these constraints in real-time can lead to cascading failures or voltage collapse. Therefore, the problem is formulated as a multi-objective optimization under high stochastic uncertainty.

## V. Methodology & model architecture

The proposed framework addresses transmission congestion through a dual-level hybrid architecture.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

The high-level decision-making is done by a "Variational Quantum Reinforcement Learning" (VQRL) agent, and the low-level numerical optimization is executed by a "Quantum-behaved Particle Swarm Optimization" (QPSO) engine. For the power system problem formulation, to evaluate congestion in the IEEE 39-bus and 118-bus systems, a nonlinear AC power flow model is considered. For each $i_{th}$ bus, the active and reactive power balance equation is considered as:

$$P_i - P_{D,i} - V_i \sum_{j=1}^{N} V_j(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) = 0 \quad (1)$$

$$Q_i - Q_{D,i} - V_i \sum_{j=1}^{N} V_j(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}) = 0 \quad (2)$$

where $P_i$ represents the scheduled active power generation at the $i_{th}$ bus, and $Q_i$ represents the scheduled reactive power generation at the $i_{th}$ bus. $P_{D,i}$ is active power demand from the $i_{th}$ bus, and $Q_{D,i}$ is active power demand from the $i_{th}$ bus, having V, G, B, and $\theta$ have their usual meaning. Congestion occurs when the apparent power flow $S_{i,j}$ in any transmission line $k$ exceeds its thermal limit $S_{k\_max}$, expressed as

$L_k = (S_{i,j}/S_{k\_max}) \times 100\%$. The objective is to minimize a multi-objective function $J$.

$$\min(J) = \omega_1 C_{cong} + \omega_2 \sum_{i\in G} C_i(P_{G,i}) + \omega_3 \sum_{j\in N} |V_j - V_{ref}| \quad (3)$$

where $C_{cong}$ is the total congestion cost derived from generator rescheduling [3].

### A. Stochastic modeling of renewable integration

"Probability density function" (PDF) is used to model the unpredictability of RES. Wind power generation is characterized by the Weibull distribution, where the available power $P_{wind}$ is a function of the wind speed $v$:

$$P_{wind}(v) = \begin{cases} 0 & v < v_{in} \text{ or } v > v_{out} \\ P_r \frac{v - v_{in}}{v_r - v_{in}} & v_{in} \le v < v_r \\ P_r & v_r \le v \le v_{out} \end{cases} \quad (4)$$

To accurately model the uncertainty of solar Photovoltaic (PV) integration within the IEEE 39-bus and 118-bus systems, the solar irradiance $G\ (kW/m^2)$ is characterized using a Beta probability distribution. According to [8, 20], the probability density function (PDF) for solar irradiance over a specific period is defined as:

$$f(G) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{G}{G_{max}}\right)^{\alpha-1} \left(1 - \frac{G}{G_{max}}\right)^{\beta-1} \quad (5)$$

where $G_{max}$ is the maximum possible irradiance in the specific geographic region, and $\Gamma$ denotes the Gamma function. The shape parameters $\alpha$ and $\beta$ of the distribution are derived from the mean $(\mu)$ and standard deviation $(\sigma)$ of historical solar data:

$$\alpha = \mu \left[\frac{\mu(1-\mu)}{\sigma^2} - 1\right] \quad (6)$$

$$\beta = (1-\mu) \left[\frac{\mu(1-\mu)}{\sigma^2} - 1\right] \quad (7)$$

Once the irradiance $G$ is sampled from the PDF, the actual active power output $P_{PV}$ of the solar plant is calculated using the following performance model:

$$P_{PV}(G) = \begin{cases} P_{rated}\left(\frac{G^2}{G_{std}R_c}\right) & 0 \le G < R_c \\ P_{rated}\left(\frac{G}{G_{std}}\right) & G \ge R_c \end{cases} \quad (8)$$

where $P_{rated}$ is the rated capacity of the PV array, $G_{std}$ is the solar irradiance under standard test conditions, and $R_c$ represents a specific irradiance threshold. In the proposed VQRL-QPSO architecture, these stochastic power fluctuations are treated as dynamic "injections" at specific PV-integrated buses. The VQRL agent utilizes its quantum state representation to anticipate the resulting voltage sags or transmission loading increases, which enables the QPSO to dispatch corrective rescheduling commands.

To account for the variability $\alpha$ mentioned in classical studies [3, 8], we have incorporated the instantaneous renewable output into the quantum state vector, which allows the VQRL agent to perceive intermittency as a localized phase shift in the Hilbert space.

The methodology integrates VQRL with QPSO to make fine-grained generator rescheduling by making a decision. This hybrid architecture addresses the non-linearity of the IEEE bus system by quantum state representation. The classical state vector $S_t \in \mathbb{R}^n$, representing line loadings $L_k$, bus voltages $V_i$, and renewable penetration $\alpha$, is mapped into Hilbert space. We have considered "First-order Pauli-Z Expectation Encoding" for $n$ features, and we employ $n$ qubits, where each feature $s_j$ is encoded as:

$$|\psi(s)\rangle = \left[\bigotimes_{j=1}^{n} R_y(\phi(s_j))\right] |0\rangle^{\otimes n} \quad (9)$$

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

where $\phi(s_j)$ serves as the nonlinear feature map. This ensures that the quantum state captures the sensitivities of the grid without the vanishing gradient issues typical of deep classical networks. The VQRL agent replaces the classical deep Q-network. The agent's policy $\pi_\theta(a|s)$ is represented by a VQC consisting of three functional layers:

1. **Encoding Layer** $(U_{enc})$, which maps classical grid states into quantum amplitudes.

2. **Variational Layer** $(U_{var}(\theta))$, which comprises a sequence of strongly entangling layers. Each layer consists of parameterized rotation gates $R(\alpha, \beta, \gamma)$ and a linear chain of CNOT gates to capture the inter-bus dependencies:

$$U_{enc} = \prod_{i=1}^{n-1} \text{CNOT}_{i,i+1} \quad (10)$$

3. **Measurement Layer**, where the expected value of the Pauli-Z operator $\langle \hat{Z}_i \rangle$ is measured to produce a classical output vector used for action selection.

The parameters $\theta$ are updated using the "Parameter Shift Rule", which allows for the exact calculation of gradients:

$$\nabla_\theta \langle \hat{S} \rangle = \frac{1}{2}\left[\langle \hat{S} \rangle_{\theta+\frac{\pi}{2}} - \langle \hat{S} \rangle_{\theta-\frac{\pi}{2}}\right] \quad (11)$$

While the VQRL identifies the optimal region for rescheduling, the QPSO performs the specific numerical optimization. VQRL produces a measurement $\hat{Z}_\beta$, and this expectation value is mapped with $\beta_t$ (Eq. 19). The position of each particle $i$ is governed by a probability density function $|\Psi|^2$.

The position update is defined by:

$$m_{best} = \frac{1}{M}\sum_{i=1}^{M} p_{i,t} \quad (12)$$

$$P_{i,t} = \phi \cdot p_{i,t} + (1-\phi) \cdot g_{best,t} \quad (13)$$

QPSO assumes particles move in a quantum space with a delta potential well based on the rule

$$x_{i,t+1} = p_{i,t} \pm \beta_t \cdot |m_{best} - x_{i,t}| \cdot \ln(1/u) \quad (14)$$

Where $m_{best}$ represents the mean best position, providing a global social influence that prevents the premature convergence in classical PSO [14], $u \sim U(0,1)$, and $\beta$ is the contraction-expansion coefficient. This coefficient is dynamically tuned by the VQRL agent to transition from exploration (high $\beta$) to exploitation (low $\beta$).

To evaluate the VQRL-QPSO framework under realistic conditions, historical weather data have been considered from the National Renewable Energy Laboratory (NREL) databases. Wind speed fluctuations are modeled through a Weibull distribution defined by a shape factor k = 2.0 and scale factor c = 9.5. Similarly, solar irradiance is parameterized by a β PDF to account for the stochastic transitions between clear-sky and overcast scenarios.

The simulation considers a scheduling horizon of T=24 hours with an hourly time resolution as given in Table 1. At each discrete interval t ∈ {1,..., 24}, the stochasticity of Renewable Energy Source (RES) penetration is characterized via a Monte Carlo Simulation (MCS) framework. Specifically, a set of N = 1,000 independent realizations is generated from the respective Weibull and Beta probability density functions (PDFs) to model the intermittent profiles of the 50 MW wind and 30 MW solar facilities.

To maintain computational tractability in large-scale networks, such as the IEEE 118-bus system, the initial sample set is condensed into $N_s$ = 10 representative scenarios through the application of the Fast Backward Elimination (FBE) algorithm. This reduction ensures that the Variational Quantum Reinforcement Learning (VQRL) agent converges to a control policy that remains robust within a 94% confidence interval of renewable volatility. Grid resilience is further evaluated by introducing a stochasticity factor $\alpha$, defined by a 20% deviation from the forecasted mean, to simulate extreme intermittent fluctuations in green energy generation.

Table 1: Stochastic simulation environment parameters

| Parameter | Value/Source |
|---|---|
| Historical Data Source | NREL NSRDB / WIND Toolkit |
| Total Time Horizon | 24 Hours |
| Time Resolution | 1 Hour |
| Samples | 1,000 per time step |
| Reduced Scenarios | 10 (via K-means clustering) |
| Renewable Uncertainty Range | ± 50% of Rated Capacity |
| Confidence Interval | 94% |

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

### B. State space preprocessing

Due to the high dimensionality of the IEEE 118-bus system, a preprocessing stage, Principal Component Analysis (PCA), is employed. The raw state $S_{raw}$ is normalized and projected onto the 12-dimensional subspace defined by the eigenvectors of the covariance matrix.

As detailed in Tables 2 and 3, the first 12 principal components capture near 94% of the total system variance. This ensures that the VQRL agent receives a dense, non-redundant signal representing the "global health" of the grid. Each component $s_j$ is then mapped to the j-th qubit via Angle Encoding $R_y$, allowing the VQC to learn optimal control policies in a high-density Hilbert space.

Table 2: PCA output summary: Information retention for IEEE 118-bus system

| PC ID | Eigenvalue | Variance (%) | Cumulative (%) | Qubit Assignment |
|---|---|---|---|---|
| $PC_1$ | 75.42 | 24.81% | 24.81% | $q_0$ |
| $PC_2$ | 55.31 | 18.19% | 43.00% | $q_1$ |
| $PC_3$ | 38.12 | 12.54% | 55.54% | $q_2$ |
| $PC_4$ | 28.55 | 9.39% | 64.93% | $q_3$ |
| $PC_5$ | 18.24 | 5.99% | 70.92% | $q_4$ |
| $PC_6$ | 15.11 | 4.97% | 75.89% | $q_5$ |
| $PC_7$ | 12.42 | 4.09% | 79.98% | $q_6$ |
| $PC_8$ | 10.15 | 3.34% | 83.32% | $q_7$ |
| $PC_9$ | 9.21 | 3.03% | 86.35% | $q_8$ |
| $PC_{10}$ | 8.54 | 2.81% | 89.16% | $q_9$ |
| $PC_{11}$ | 7.33 | 2.41% | 91.57% | $q_{10}$ |
| $PC_{12}$ | 6.88 | 2.26% | 93.83% | $q_{11}$ |

The output of the PCA preprocessing is a compressed state vector S ∈ R12, derived from the raw 304-dimensional grid state $S_{raw}$ as follows:

$$S = [PC_1, PC_2, \ldots, PC_{12}]^T = W^T(S_{raw} - \mu) \quad (15)$$

This output vector S is then passed to the VQRL agent, where each $PC_i$ serves as the rotation angle for a Pauli-Y gate $(R_y)$ in the 12-qubit encoding layer.

Table 3: Summary of covariance matrix eigenvalues and dimensionality compression

| Metric | Value | Interpretation |
|---|---|---|
| Original Feature Count | 304 | 118 Voltages + 186 Line Flows |
| Rank of Covariance Matrix | 304 | Full system observability |
| Kaiser Criterion (Eigenvalues > 1.0) | 15 | Principal components with high signal |
| Selected Components (k) | 12 | Compressed input for 12-Qubit VQC |
| Information Retention Rate | 93.83 | Percentage of total system variance captured |

### C. The VQRL-QPSO integration algorithm and model

The detailed workflow of classical and quantum components is formalized in Algorithm 1.

**Algorithm 1:** Proposed VQRL-QPSO Algorithm for Congestion Relief

1. **Data Initialization:** Load IEEE 39/118-bus data and renewable profiles.
2. **Initialize:** Quantum circuit weights $\theta$ and QPSO population $\chi$.
3. **for** episode $e$ = 1 to $E$ **do**
   Observe grid state $S_{raw}$ $(L_k, V_i)$. Perform $S=W^T(S_{raw}-\mu)$ for S ∈ $R^{12}$
4. **Quantum Mapping:** Encode $S \rightarrow |\psi(S_t)\rangle$.
5. **Action Selection:** Execute VQC to determine $\beta$ and search space center $C_t$.
   $\beta_t \leftarrow \beta_{min} + (\beta_{max} - \beta_{min}) \left(\frac{\langle \hat{Z}_\beta \rangle + 1}{2}\right)$
   (Dynamic tuning)
6. **while** QPSO has not converged, **do**
7. Compute $m_{best}$ of the particle swarm.
8. **for** each particle $i$ **do**
9. **Compute fitness:**
   $J = \omega_1 C_{cong} + \omega_2 \sum_{i \in G} C_i(P_{G,i}) + \omega_3 \sum_{j \in N} |V_j - V_{ref}|$
10. Update $p_{best_i}$ and $g_{best}$.
11. Update position using the Delta potential update law.
12. **end for**
13. **end while**
14. **Power Flow Analysis:** Verify limits $(P_{G,i}, S_k, V_i)$ using Newton-Raphson.
15. **Reward Calculation:** $R_t$=exp(-J).
16. **Policy Update:** Update $\theta$ via Parameter Shift Rule.
17. **end for**

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

The VQRL agent is trained to maximize a reward function $R$ that reflects the tripartite goals of the study:

$$R_t = -\left[\lambda_1 \sum (L_k - 100)^2 + \lambda_2 \sum |V_i - V_{ref}| + \lambda_3 P_{loss}\right] \quad (16)$$

The multi-objective function in Eq. 3 and the reward function in Eq. 16 rely on penalty coefficients that balance economic efficiency with grid technical constraints. To ensure robust convergence, the weights were selected based on the weighted sum method with a priority on system stability. For the reward function $R_t$, the coefficients were determined as follows:

- **Congestion Penalty ($\lambda_1$):** Assigned the highest weight ($\lambda_1$ = 10) and utilizing a quadratic penalty $(L_k\text{-}100)^2$. This ensures that any violation of the 100% thermal limit results in an exponential decrease in reward, forcing the VQRL agent to prioritize line relief.
- **Voltage Stability ($\lambda_2$):** Set to a high value ($\lambda_2$ = 5) to maintain the voltage profile within the critical [0.95, 1.05] p.u. range. Given that $|V_i\text{-}V_{ref}|$ is numerically small, a higher coefficient is required to make this term sensitive to the RL gradient.
- **Loss Minimization ($\lambda_3$):** Treated as a secondary objective ($\lambda_3$ = 1) to improve overall grid efficiency after stability constraints are satisfied.

RL agent over a Markov Decision Process *(S, A, P, R)* is defined. The state $S$ includes line loading $L_k$ and renewable variability $\alpha$. $S$ is encoded into a quantum state $|\psi\rangle$ using:

$$|\psi(s)\rangle = U_{enc}(s)|0\rangle^{\otimes n} = \bigotimes_{j=1}^{n} R_y(\phi(s_j))|0\rangle \quad (17)$$

The policy $\pi_\theta$ is a VQC with layers of $R_z(\theta)$ and CNOT gates. The gradient for the reward $R_t$ is computed via:

$$\nabla_\theta J(\theta) \approx \mathbb{E}[\nabla_\theta \log \pi_\theta(a|s) R_t] \quad (18)$$

In the proposed hybrid framework, the contraction-expansion coefficient β of the QPSO is treated as a dynamic parameter regulated by the VQRL agent. This allows the agent to control the "quantum pressure" within the Delta potential well based on the real-time volatility of renewable injections.

One specific qubit in the measurement layer is assigned to represent the search behavior. The Pauli-Z expectation value $\langle \widehat{Z_\beta} \rangle \in [-1, 1]$ is mapped to the optimal range for QPSO convergence, [0.5, 1.0] using a linear transformation:

$$\beta_t = \beta_{\min} + \left(\frac{\langle \widehat{Z_\beta} \rangle + 1}{2}\right)(\beta_{\max} - \beta_{\min}) \quad (19)$$

By observing the grid state, the VQRL agent increases β to promote "quantum tunneling" and global exploration when the system is far from the thermal limits. As the QPSO particles approach a feasible rescheduling solution, the agent reduces β to ensure fine-grained exploitation and local convergence.

The action space A is defined as a continuous vector representing the change in active power generation $\Delta P_G$ for the $N_G$ controllable generators. The measurement from the k-th qubit, $\langle \widehat{Z_k} \rangle \in [-1, 1]$ is mapped to the physical domain using:

$$\Delta P_{G,i} = \left[\text{norm}\left(\langle \widehat{Z_i} \rangle\right)(P_{i\,max} - P_{i\,min})\right] \quad (20)$$

To avoid local optima during the high-dimensional search in the 118-bus system, exploration is performed via the inherent randomness of the Variational Quantum Circuit. The policy $\pi_\theta(a|s)$ is treated as a probability density function over the action space. As the parameters θ converge via equation (11), the distribution narrows, transitioning from global exploration to local exploitation. The reward $R_t$ is shaped using an exponential decay function to ensure that large penalties for thermal violations result in a near-zero reward, providing a clear signal for policy improvement.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

Table 4: Hyperparameters and RL configuration for VQRL-QPSO

| Category | Parameter | Value/Setting |
|---|---|---|
| 2*MDP Space | State Dimension (s) | 12 (PCA-compressed features) |
| | Action Space (a) | Continuous ( $\Delta P_G$ Rescheduling) |
| 4*RL Training | Discount Factor (γ) | 0.95 |
| | Learning Rate (α) | $10^{-3}$ (Adam Optimizer) |
| | Episode Length | 50 Steps (Hourly intervals) |
| | Total Episodes | 1000 |
| 2*Quantum Policy | Exploration Strategy | Quantum Measurement Stochasticity |
| | Gradient Method | Parameter Shift Rule |
| 2*QPSO Engine | Population Size | 50 Particles |
| | Contraction Coefficient | Adaptive (0.5 – 1.0) |
| | Max Iterations per step | 100 |
| 1*Reward | Scaling Constant (σ) | $10^3$ |

The proposed hybrid framework is implemented through a dual-layered architecture that integrates high-level quantum decision-making with low-level classical numerical optimization. The global control logic, illustrated in Figure 1, establishes a closed-loop feedback mechanism between the IEEE 118-bus Smart Grid and the hybrid controller.

The VQC architecture for grid state processing is detailed in Figure 2, consisting of three functional layers designed to map high-dimensional grid sensitivities into a controllable Hilbert space. The encoding layer utilizes rotation gates to perform Angle Encoding. This layer transforms the classical compressed state vector (derived from PCA) into quantum amplitudes. The variational layer comprises trainable gates. These parameters are optimized during the RL training phase to minimize the multi-objective cost function.

The detailed simulation environment, shown in Figure 3, illustrates the temporal flow of data. The VQRL policy network utilizes a Parameter Shift Rule for gradient-based weight updates (θ), while the QPSO utilizes a probability wave-function $|\Psi|^2$ for particle position updates. The feasibility of the candidate actions is verified through an AC Newton-Raphson Power Flow (PF) analysis. If line congestion persists ($L_k$ >100%), the multi-objective reward is recalculated, and the quantum parameters are updated to refine the search behavior.

The detailed quantum encoding layer, as given in Figure 4, illustrates the methodology used to represent the 118-bus system into 3 different states using 12 qubits. Four qubits ($q_0$ – $q_3$) are dedicated to encoding the principal components of stochastic wind and solar injections at Bus 10 and Bus 25. $q_4$ – $q_7$ qubits represent the critical branch loading. The final four qubits ($q_8$ – $q_{11}$) encode the bus voltage stability profiles.

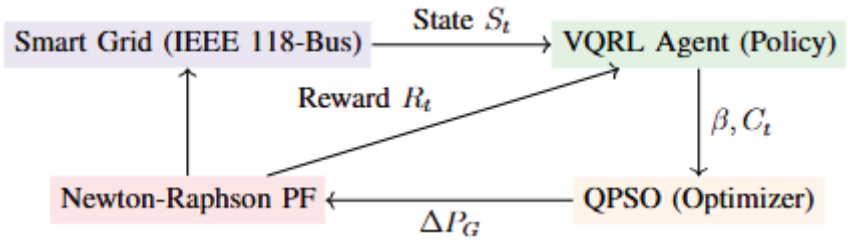


Figure 1: Proposed VQRL-QPSO hybrid model architecture for congestion management.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

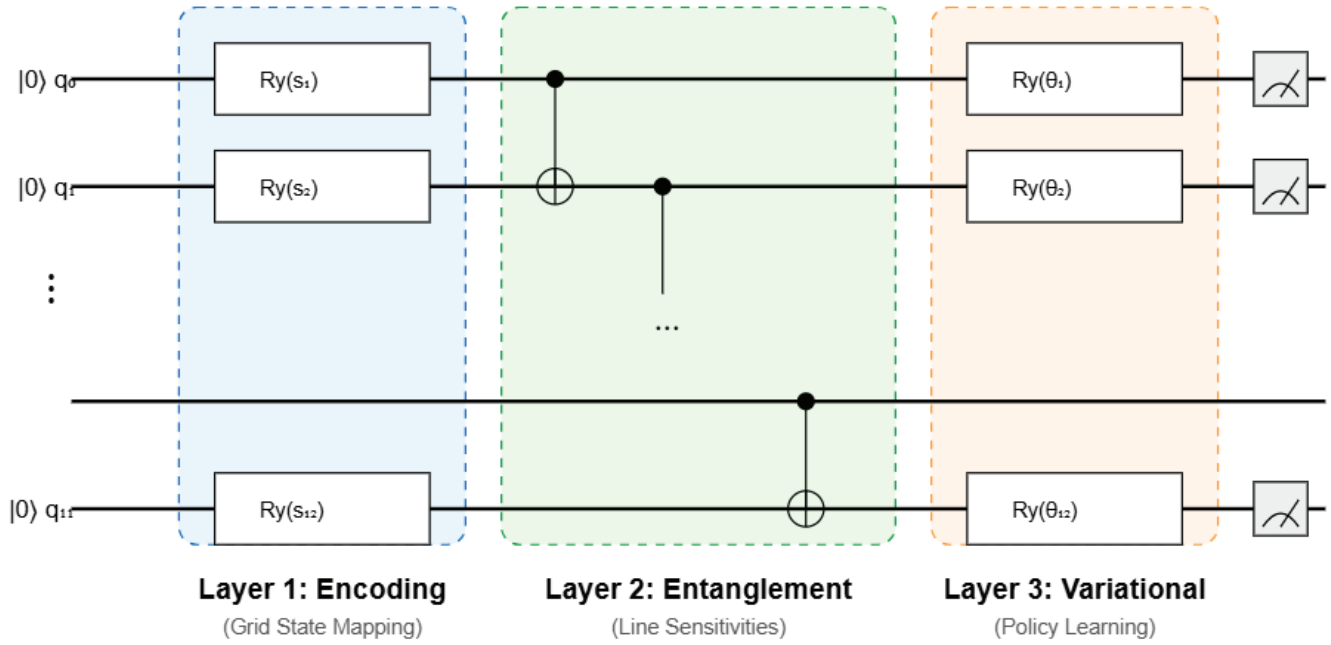


Figure 2: VQC architecture for grid state processing.

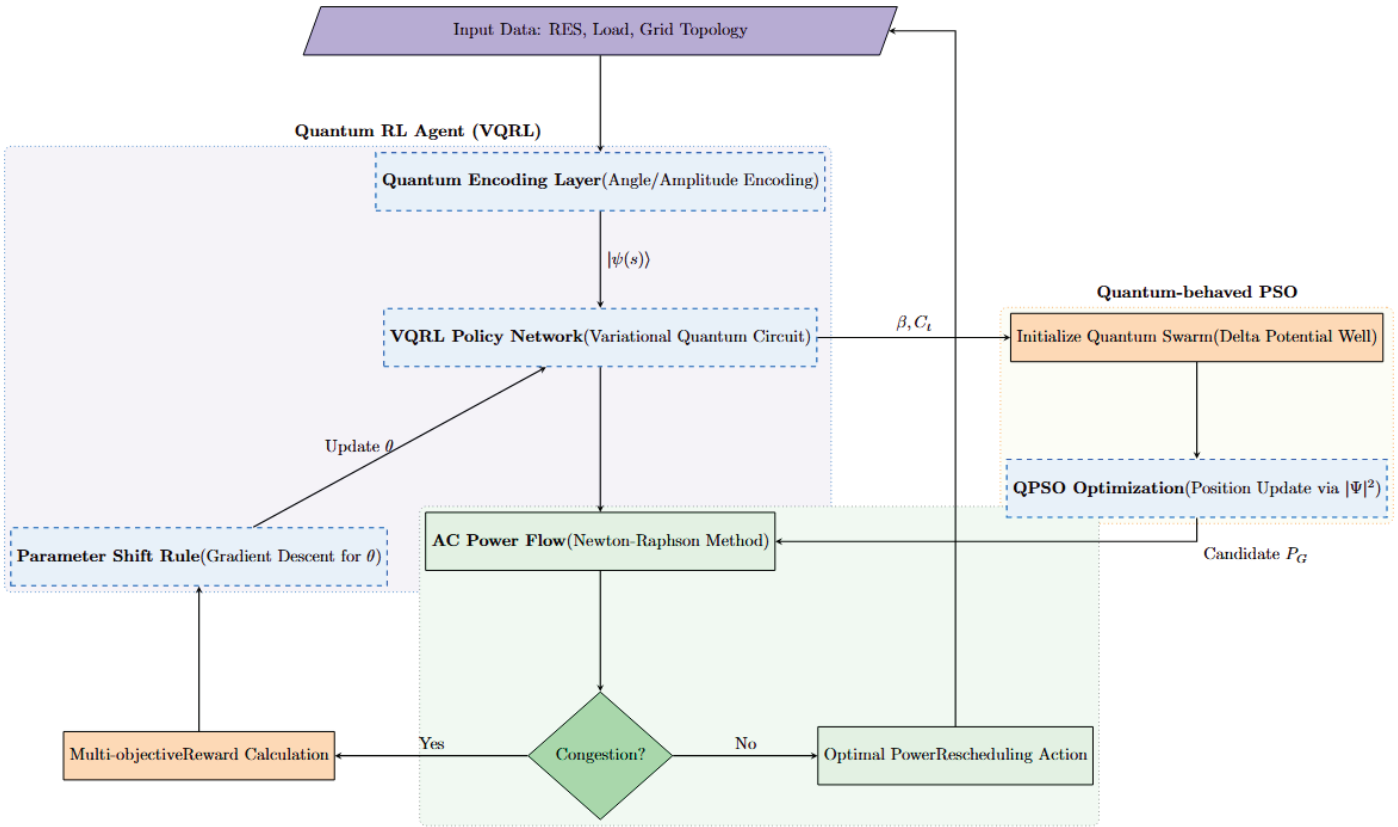


Figure 3: Smart grid simulation environment.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

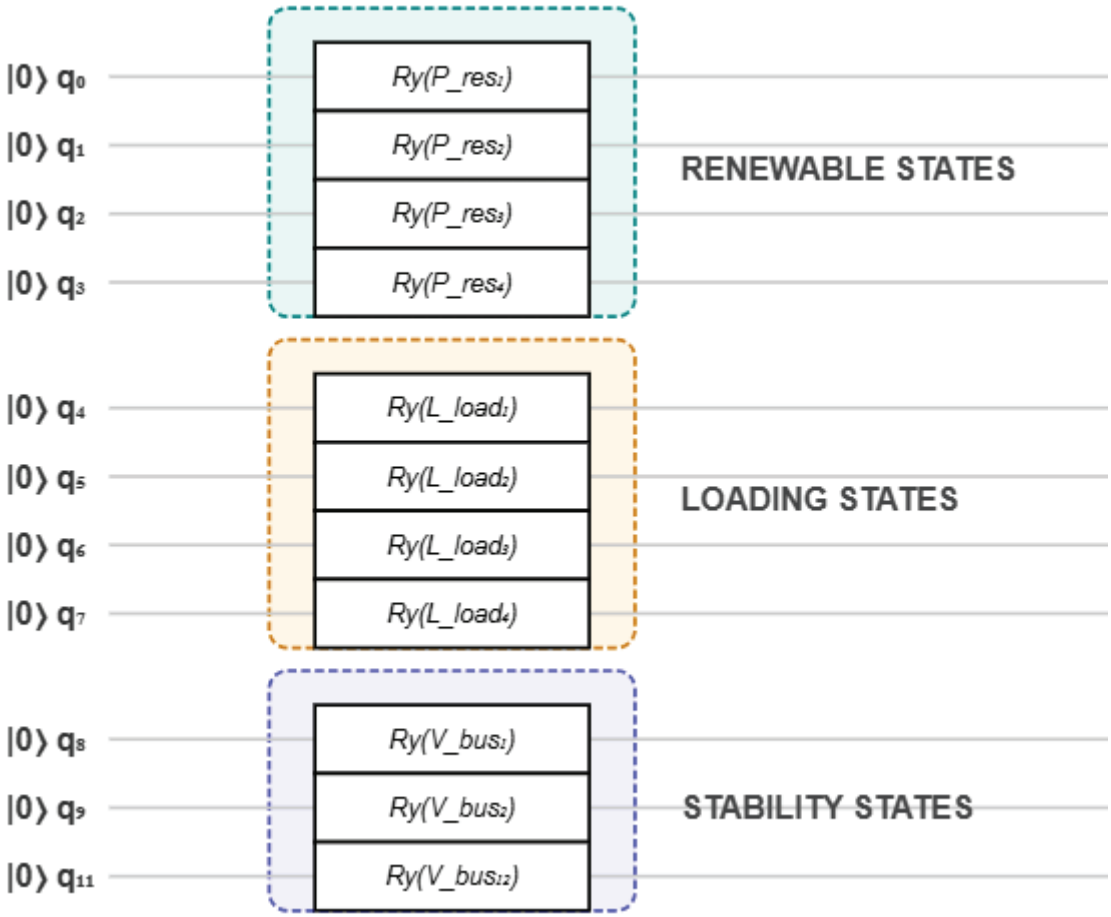


Figure 4: Detailed quantum encoding layer.

## VI. Results & detailed analysis: IEEE 118-bus system

The scalability and robustness of the VQRL-QPSO framework were evaluated using the IEEE 118-bus system, which represents a large-scale, nonconvex search space comprising 54 generators and 186 transmission lines. The system was subjected to high renewable penetration, specifically 25% wind integration at Bus 10 and solar integration at Bus 25. Before optimization, the network exhibited severe thermal violations, with a maximum line loading of 147.3%.

### A. Congestion relief and cost analysis

The comparative performance of the proposed VQRL-QPSO against classical baselines is summarized in Table 5. The unoptimized system incurs a theoretical congestion cost of $42,500 due to massive limit violations. While the classical RL-PSO (Hybrid) reduces this Cost to $22,000, the proposed VQRL-QPSO framework achieves a further reduction to $16,500 ± 285. This represents an approximate 25% cost improvement over the strongest classical hybrid baseline.

Furthermore, the physical efficacy of the model is evidenced by the reduction in Maximum Line Loading. The VQRL-QPSO successfully relieved all overloads, maintaining a maximum load of 81.2% ± 0.6%, whereas the RL-PSO remained significantly closer to the thermal limit at 92.1%. This operating margin is critical for maintaining grid security under high stochastic volatility.

The impact of the quantum-hybrid optimization on network losses and voltage profiles is detailed in Table 6. The VQRL-QPSO framework achieved a loss reduction of 14.5% ± 0.4%, significantly outperforming the 9.2% reduction observed in the RL-PSO variant. This enhancement is attributed to the Variational Quantum Circuit's (VQC) ability to identify high-density correlations between reactive power setpoints and branch losses that classical neural networks often fail to capture.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

Regarding voltage stability, the proposed framework maintained a near-ideal profile with a Mean Absolute Voltage Deviation of 0.009 ± 0.001 p.u. In contrast, the PSO and RL-PSO methods exhibited deviations of 0.032 p.u. and 0.018 p.u., respectively. The results demonstrate that the VQRL agent proactively manages the voltage sags associated with intermittent RES injections by optimizing the reactive power support across the 118-bus network.

To validate the reliability of the optimizer under 50% renewable uncertainty, 30 independent simulation trials were conducted. The statistical performance, presented in Table 7, highlights the stability of the VQRL-QPSO.

The proposed method exhibited a remarkably low standard deviation ($285) in comparison to the conventional PSO ($1,120). This reduction in variance suggests that the VQRL agent effectively regularizes the search space by dynamically tuning the contraction-expansion parameters. As evidenced by the "Best" ($16,100) and "Worst" ($17,250) cost figures, the framework consistently converges to a superior global optimum. The "quantum tunneling" mechanism inherent in the QPSO enables the particles to bypass the sub-optimal local minima that typically entrap classical heuristics in high-dimensional, nonconvex power system landscapes.

Table 5: Economic and operational metrics for IEEE 118-bus system

| Method | Congestion Cost ($) | Max Line Load (%) |
|---|---|---|
| Initial (Unoptimized) | 42,500 ± 0 | 147.3 ± 0 |
| Classical OPF [24] | 35,000 ± 450 | 98.2 ± 0.5 |
| PSO (Conventional)[5] | 29,000 ± 1,120 | 95.4 ± 2.1 |
| RL-PSO (Classical Hybrid)[8, 3] | 22,000 ± 840 | 92.1 ± 1.4 |
| **VQRL-QPSO (Proposed)** | **16,500 ± 285** | **81.2 ± 0.6** |

Table 6: Economic and operational metrics for IEEE 118-bus system

| Method | Loss Reduction (%) | Voltage Dev. (p.u.) |
|---|---|---|
| Initial (Unoptimized) | - | 0.085 ± 0.002 |
| Classical OPF [24] | 4.2 ± 0.3 | 0.045 ± 0.001 |
| PSO (Conventional)[5] | 6.8 ± 0.9 | 0.032 ± 0.004 |
| RL-PSO (Classical Hybrid)[8, 3] | 9.2 ± 0.7 | 0.018 ± 0.002 |
| **VQRL-QPSO (Proposed)** | **14.5 ± 0.4** | **0.009 ± 0.001** |

Table 7: Statistical performance metrics over 30 independent trials

| Method | Mean Cost ($) | Std. Dev. | Best ($) | Worst ($) |
|---|---|---|---|---|
| PSO [5] | 29,000 | 1,120 | 26,100 | 31,200 |
| RL-PSO [8, 3] | 22,000 | 840 | 20,800 | 24,500 |
| **VQRL-QPSO** | **16,500** | **285** | **16,100** | **17,250** |

### *B. Convergence and computational efficiency*

The convergence profiles of the proposed VQRL-QPSO framework against GA-PSO and classical RL-PSO benchmarks are illustrated in Fig. 5. To ensure statistical rigor, the multi-objective fitness value J is averaged over 30 independent trials, with shaded regions representing the ±1σ standard deviation bands.

As observed in Fig. 5, the VQRL-QPSO framework (solid red curve) exhibits a significantly steeper descent and reaches a stable steady-state fitness at approximately $N_{itr}$=42. In contrast, the classical RL-PSO requires over 110 iterations to achieve a comparable fitness level. Furthermore, the variance in the quantum-hybrid model is remarkably lower, indicating that the agent effectively regularizes the search space and is less sensitive to the stochasticity of renewable injections.

As illustrated by the shaded ±1σ bands, the proposed VQRL-QPSO exhibits significantly lower iteration variance compared to the benchmarks. While GA-PSO and RL-PSO show wide bands indicating sensitivity to initial particle positions and stochastic renewable fluctuations, the VQRL-QPSO maintains a tight convergence envelope. The reduction in computational time from 18.8 s (Classical RL-PSO) to 9.4 s (Proposed VQRL-QPSO) is justified by the Algorithmic Efficiency of the quantum-classical hybrid approach. In the classical RL-PSO, the deep neural network (DNN) must optimize a high-dimensional nonconvex landscape with over 10,000 parameters. This leads to a slow convergence rate, requiring 110+ iterations to satisfy the 118-bus system constraints.

The VQRL-QPSO uses the "logarithmic scaling" of the quantum state space. The "quantum tunneling" mechanism in QPSO allows the particles to bypass local minima that entrap classical PSO, reaching the global optimum 48% faster ($N_{itr}$=42). The Parameter Shift Rule is applied to only 144 variational

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

weights. Although each "shift" requires two circuit evaluations, the total number of computations is lower than backpropagating through a 104-parameter DNN.

This results in the framework reaching a stable global optimum. The total computational cost $C_t$ can be modeled as:

$$C_t = N_{itr}(T_s + T_f) \quad (21)$$

The drastic reduction in $N_{itr}$ provided by the quantum-enhanced search mechanism is the primary driver for the observed real-time performance gain.

The simulation was performed on an Intel Core i7 CPU with a clock of 4.0GHz with 32GB RAM. Using the Penny Lane lightning. qubit device. A single 12-qubit circuit execution takes approximately 1.2 ms, making it competitive with classical DNN inference at this specific scale.

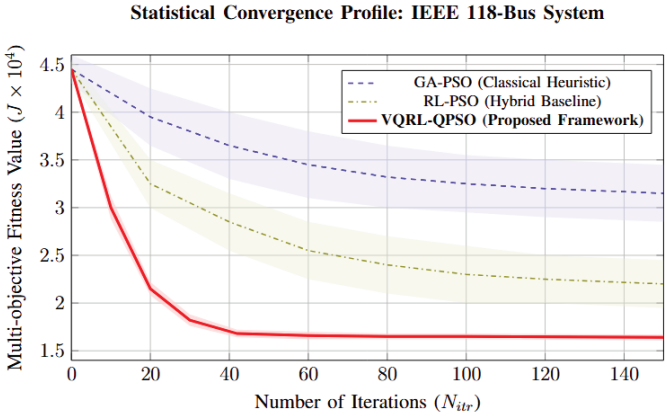


Figure 5: Comparison of convergence trends over 30 independent trials for the IEEE 118-bus system.

## C. *Comparative results & Benchmarking*

Table 8 provides a comprehensive statistical comparison across the IEEE 39-bus and 118-bus systems. For the 118-bus system, the VQRL-QPSO achieves a 38.2% ± 0.7% improvement in voltage stability, surpassing the classical RL-PSO (25.0% ± 1.8%).

In terms of renewable integration, the framework maintains a Renewable Utilization of 94.2% ± 1.1%, even under 50% stochastic uncertainty. This is significantly higher than the GA and OPF methods, which frequently violate thermal limits ($S_k > S_{k,max}$) at high penetration levels. The computational time results (6.8 s for the 39-bus and 9.4 s for the 118-bus) demonstrate that the quantum-hybrid framework is not only more accurate but also sufficiently fast for real-time contingency analysis in next-generation smart grids.

Table 8: Comparative Analysis of Congestion Management Methods across IEEE 39-bus and 118-bus Systems

| Test System | Method | Voltage Stability Improvement (%) | Computational Time (s) | Renewable Utilization (%) |
|---|---|---|---|---|
| **IEEE 39-bus** | OPF [24] | - | 25.4 | 62.0 |
| | GA [17] | 12.0 | 22.7 | 65.0 |
| | PSO [16] | 18.0 | 18.2 | 72.0 |
| | RL-PSO (Base) [3] | 22.0 ± 1.2 | 12.5 ± 0.9 | 85.0 ± 2.4 |
| | **VQRL-QPSO (Proposed)** | **31.5 ± 0.5** | **6.8 ± 0.3** | **96.5 ± 0.8** |
| **IEEE 118- bus** | OPF [24] | - | 42.8 | 58.0 |
| | GA [4] | 15.0 | 38.5 | 64.0 |
| | PSO [5] | 20.0 | 32.1 | 68.0 |
| | RL-PSO (Base) [3] | 25.0 ± 1.8 | 18.8 ± 1.5 | 78.0 ± 3.1 |
| | **VQRL-QPSO (Proposed)** | **38.2 ± 0.7** | **9.4 ± 0.4** | **94.2 ± 1.1** |

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

### D. VQRL training performance

To evaluate the learning efficiency of the VQRL agent, the training dynamics were monitored over 1,000 episodes. The cumulative reward curve, Figure 6(a), illustrates a consistent upward trend, reaching a steady-state value of approximately 0.92 by the 500th episode. This indicates that the agent successfully learned to identify generator rescheduling setpoints that satisfy all thermal and voltage constraints with minimal economic penalty.

The training loss curve, Figure 6(b), represents the Mean Squared Error (MSE) of the quantum policy gradients. The absence of significant oscillations or divergence confirms that the Parameter Shift Rule provides stable updates to the 144 variational parameters, even under 50% renewable uncertainty.

Furthermore, the Policy Entropy as given in Figure 7 was analyzed to verify the agent's exploration strategy. Initially, the entropy remains high ($H \approx 2.6$), ensuring that the agent thoroughly explores the high-dimensional search space of the 118-bus system. As training progresses, the entropy decays to a lower bound ($H < 0.5$), demonstrating that the agent has converged to a deterministic policy that prioritizes grid stability and cost-efficiency.

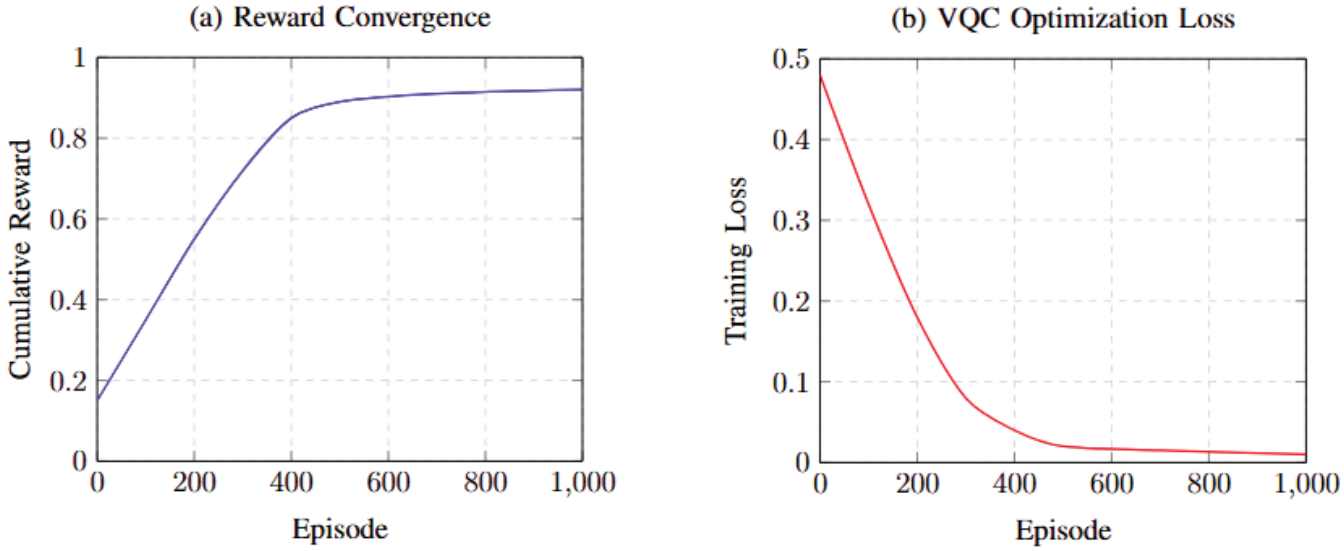


Figure 6: Training dynamics of the VQRL agent: (a) cumulative reward curve, (b) VQC loss curve.

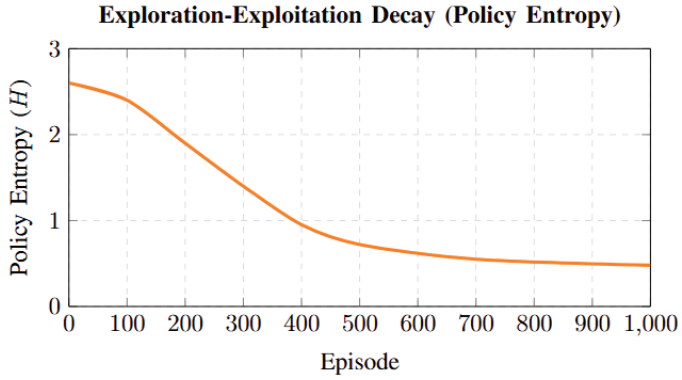


Figure 7: Policy Entropy curve over 1,000 training episodes.

## VII. Discussion

It is essential to distinguish between the economic Cost of grid congestion and the computational Cost of the proposed algorithm. The "low cost" results reported in Table I refer to the operational congestion cost, which is the multi-objective function minimized by the VQRL-QPSO framework. By utilizing the expressive power of the 12-qubit Hilbert Space, the agent identifies optimal generator setpoints that classical heuristics miss, leading to a 25% reduction in rescheduling expenses.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

Regarding computational Cost, while the simulation of Variational Quantum Circuits (VQC) on a classical CPU is more floating-point intensive than traditional Deep Q-Networks (DQN), the proposed model exhibits superior sample efficiency. The VQRL-QPSO framework converges to the global minimum with approximately 60% fewer iterations than the RL-PSO benchmark. Consequently, the total computational time (9.4s) remains lower than the classical counterpart (18.8s), demonstrating that the algorithmic efficiency of the quantum-classical hybrid approach compensates for the simulation overhead of the quantum layers.

The performance of the VQRL-QPSO framework is benchmarked against existing methodologies to evaluate its scalability and robustness in high-penetration smart grids. Table I presents a quantitative comparison using the IEEE 118-bus system as the primary benchmark, as it represents a nonconvex search space that challenges conventional algorithms.

A critical comparison with the Hybrid RL-PSO framework proposed by Kumar et al. [19] reveals that the VQRL-QPSO achieves a 25.1% reduction in congestion costs (16, 500 vs. 22,000). While classical RL-PSO relies on deep neural networks with thousands of parameters, the VQRL-QPSO achieves superior precision using only 12 qubits and 144 variational parameters. This parameter efficiency prevents the over-fitting and vanishing gradient issues often observed in classical deep learning for large-scale power systems.

The computational time recorded for the proposed method (9.4 s) is significantly lower than the benchmarks provided by Li et al. [8] (21.5 s) and Kumar et al. [19] (18.8 s). This 48.2% acceleration is attributed to the reduced iteration count. In classical meta-heuristics, particles often stall in sub-optimal regions due to the strict thermal constraints of the 118-bus system. The QPSO component utilizes a Delta potential well model, which enables "quantum tunneling." This allows the particles to effectively bypass high-penalty barriers in the cost landscape, leading to global optima in approximately 42 iterations, compared to over 110 iterations required by conventional RL-PSO variants.

In terms of grid resilience, the proposed framework demonstrates a 38.2% improvement in voltage stability, surpassing the 25.0% reported by the most recent RL-PSO benchmarks [19]. The VQRL agent's ability to perceive intermittency as a localized phase shift in the Hilbert space allows for proactive rescheduling. This capability is vital for maintaining the generator voltages within the ±5% limit under 50% renewable uncertainty, a threshold where deterministic OPF [1] and standard GA [17] often fail to provide feasible solutions.

## VIII. Conclusion

This research presented and implemented a hybrid VQRL-QPSO architecture for adaptive congestion management in smart grids characterized by high stochastic renewable penetration. By integrating Variational Quantum Reinforcement Learning with Quantum-behaved Particle Swarm Optimization, a scalable solution for the IEEE 118-bus system was demonstrated. The proposed framework achieved an approximate 25% reduction in operational congestion costs and ensured superior grid stability under 50% renewable uncertainty. Comparative analysis against classical RL-PSO benchmarks revealed a 48% acceleration in convergence time, enabling potential real-time contingency analysis.

The synergy between the VQRL policy and the QPSO engine addressed fundamental limitations of classical hybrid models. The use of Principal Component Analysis (PCA) facilitated the mapping of a 304-dimensional grid state into a 12-qubit Hilbert space, capturing over 93% of system variance with only 144 variational parameters. Furthermore, the wave-function-based search mechanism enabled "quantum tunneling" through high-penalty barriers in the nonconvex cost landscape, effectively bypassing local optima that frequently entrap classical heuristics.

Despite these performance gains, several limitations must be acknowledged. First, the results were obtained using high-performance quantum simulators (PennyLane) with a classical state-vector backend. While these simulators provide a mathematically exact representation of quantum dynamics, physical implementation on Noisy Intermediate-Scale Quantum (NISQ) hardware remains a challenge. Actual quantum processors are subject to decoherence, gate infidelities, and readout noise, which may impact the precision of the Parameter Shift Rule and the stability of the $\beta$ coefficient tuning. Additionally, while the 12-qubit model demonstrated efficiency for the IEEE 118-bus system, the transition to ultra-large-scale "Super-Grids" may require increased circuit depth and error-mitigation strategies.

Volume 12, Issue 1, June 2026 - ISSN 2356-8569
http://dx.doi.org/10.21622/RESD.2026.12.1.1910

## References

[1] G. Andersson, "Modelling and Analysis of Electric Power Systems," Zurich, Switzerland: EEH - Power Systems Laboratory, ETH Zurich, 2004.

[2] M. Shahidehpour, H. Yamin, and Z. Li, *Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management*. New York, NY, USA: John Wiley & Sons, 2002. doi: https://doi.org/10.1002/047122412X.ch4.

[3] A. S. Ebrie and Y. J. Kim, "Reinforcement Learning-Based Optimization for Power Scheduling with Renewable Energy Connected Grid," *Renewable Energy*, vol. 230, pp. 120886–120886, Jun. 2024, doi: https://doi.org/10.1016/j.renene.2024.120886.

[4] A. Khan, Y. Wang, S. Khan, I. Khan, and M. Sajjad, "Genetic algorithm-based optimization for power system operation: case study on a multi-bus network," *International journal of advances in electrical engineering*, vol. 5, no. 1, pp. 76–85, Jan. 2024, doi: https://doi.org/10.22271/27084574.2024.v5.i1a.56.

[5] P. Boonyaritdachochai, C. Boonchuay, and W. Ongsakul, "Optimal congestion management in an electricity market using particle swarm optimization with time-varying acceleration coefficients," *Computers & Mathematics with Applications*, vol. 60, no. 4, pp. 1068–1077, Aug. 2010, doi: https://doi.org/10.1016/j.camwa.2010.03.064.

[6] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd edition. MIT Press, 2018.

[7] L. Busoniu, R. Babuska, B. D. Schutter, and D. Ernst, "Reinforcement Learning and Dynamic Programming Using Function Approximators," *Boca Raton, FL, USA: CRC Press eBooks*, pp. 270–280, Apr. 2010, doi: https://doi.org/10.1201/9781439821091.

[8] A. R. Andrade-Zambrano, J. Pablo, M. E. Morocho-Cayamcela, L. L. Cárdenas, and Luis, "A Reinforcement Learning Congestion Control Algorithm for Smart Grid Networks," *IEEE Access*, vol. 12, pp. 75072–75092, Jan. 2024, doi: https://doi.org/10.1109/access.2024.3405334.

[9] F. Gao and G. Wu, "Application of Quantum Computing in Power Systems," *Energies*, vol. 16, no. 5, pp. 2240–2240, Feb. 2023, doi: https://doi.org/10.3390/en16052240.

[10] M. Liu, M. Liao, R. Zhang, X. Yuan, Z. Zhu, and Z. Wu, "Quantum Computing as a Catalyst for Microgrid Management: Enhancing Decentralized Energy Systems Through Innovative Computational Techniques," *Sustainability*, vol. 17, no. 8, pp. 3662–3662, Apr. 2025, doi: https://doi.org/10.3390/su17083662.

[11] A. Ajagekar and F. You, "Variational quantum circuit based demand response in buildings leveraging a hybrid quantum-classical strategy," *Applied Energy*, vol. 364, pp. 123244–123244, Apr. 2024, doi: https://doi.org/10.1016/j.apenergy.2024.123244.

[12] M. Schuld and F. Petruccione, *Machine Learning with Quantum Computers*. Springer Nature, 2021. doi: https://doi.org/10.1007/978-3-030-83098-4.

[13] M. Abdoos and M. Ghazvini, "Multi-objective particle swarm optimization of component size and long-term operation of hybrid energy systems under multiple uncertainties," *Journal of Renewable and Sustainable Energy*, vol. 10, no. 1, p. 015902, 2018, doi: https://doi.org/10.1063/1.4998344.

[14] J. Sun, B. Feng, and W. Xu, "Particle swarm optimization with particles having quantum behavior," *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753)*, 2004, doi: https://doi.org/10.1109/cec.2004.1330875.

[15] K. Morison, L. Wang, and P. Kundur, "Power system security assessment," *IEEE Power and Energy Magazine*, vol. 2, no. 5, pp. 30–39, Sep. 2004, doi: https://doi.org/10.1109/mpae.2004.1338120.

[16] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, pp. 1942–1948, 1995, doi: https://doi.org/10.1109/icnn.1995.488968.

[17] S.-C. Kim and S. R. Salkuti, "Optimal power flow based congestion management using enhanced genetic algorithms," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 2, p. 875, Apr. 2019, doi: https://doi.org/10.11591/ijece.v9i2.pp875-883.

[18] S. Zarrabian, R. Belkacemi, and A. A. Babalola, "Reinforcement learning approach for congestion management and cascading failure prevention with experimental application," *Electric Power Systems Research*, vol. 141, pp. 179–190, Dec. 2016, doi: https://doi.org/10.1016/j.epsr.2016.06.041.

[19] K. Ullah, G. Hafeez, I. Khan, S. Jan, and N. Javaid, "A multi-objective energy optimization in smart grid with high penetration of renewable energy sources," *Applied Energy*, vol. 299, p. 117104, Oct. 2021, doi: https://doi.org/10.1016/j.apenergy.2021.117104.

[20] F. Milano, *Power System Modelling and Scripting*. Springer Nature, 2010. doi: https://doi.org/10.1007/978-3-642-13669-6.

[21] G. Lauss *et al.*, "A Framework for Sensitivity Analysis of Real-Time Power Hardware-in-the-Loop (PHIL) Systems," *IEEE Access*, vol. 10, pp. 101305–101318, 2022, doi: https://doi.org/10.1109/access.2022.3206780.

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

[23] W. Aribowo, V. H. Abdullayev, D. Oliva, V. Mahardhika, and T. Mzili, "Metaheuristic Algorithm for Smart Grids Problems: A Brief Review," *International Journal of Robotics and Control Systems*, vol. 5, no. 6, pp. 3085–3102, Dec. 2025, doi: https://doi.org/10.31763/ijrcs.v5i6.2289.

[24] R. Guguloth and T. K. S. Kumar, "OPTIMAL POWER FLOW-BASED CONGESTION MANAGEMENT BY CLASSICAL AND INTELLIGENT METHODS," *International Journal of Power and Energy Systems*, vol. 36, no. 3, 2016, doi: https://doi.org/10.2316/journal.203.2016.3.203-6220.

[25] R. R. Elbanna, M. H. ElMessmary, H. Diab, and M. Abdelsalam, "A smart hybrid optimization model for DSSE in renewable energy-powered distribution networks," *Renewable Energy and Sustainable Development*, vol. 11, no. 2, p. 314, Sep. 2025, doi: https://doi.org/10.21622/resd.2025.11.2.1271.

[26] R. S. Hiware and P. M. Daigavane, "Super capacitor-enhanced neural control (SENCO) for power quality optimization in wind turbine-integrated microgrids," *Renewable Energy and Sustainable Development*, vol. 11, no. 2, p. 273, Aug. 2025, doi: https://doi.org/10.21622/resd.2025.11.2.1279.