High-Accuracy and Robust SOC Estimation Using a PSO-Optimized AFCM–RRBF Hybrid Model

Donya Souidi
Université de Tunis, ENSIT
Laboratoire LISIER
Tunis, Tunisie
Donyasouidi123@gmail.com

Jaouher Chrouta
Université de Tunis, ENSIT
Laboratoire LISIER
Tunis, Tunisie
chroutajawher@yahoo.fr

Achref Jabeur Telmoudi
Université de Tunis, ENSIT
Laboratoire LISIER
Tunis, Tunisie
achraf.telmoudi@ieee.org

***Abstract*—Accurate estimation of the state of charge (SOC) is a fundamental requirement for ensuring both the safety and optimal performance of lithium-ion battery management systems, particularly under highly nonlinear and time-varying operating conditions. Nevertheless, traditional estimation techniques remain limited by cumulative drift and a pronounced sensitivity to measurement noise, which can significantly degrade their reliability in practical scenarios. To overcome these challenges, this study introduces a hybrid modeling framework that integrates adaptive fuzzy clustering with a recurrent radial basis function (RRBF) neural network. In addition, a particle swarm optimization (PSO) strategy is employed to achieve global tuning of the model parameters, thereby enhancing convergence and generalization capabilities. The effectiveness of the proposed approach is validated through comprehensive experimental evaluations. The results indicate a high level of estimation accuracy, with root mean square error (RMSE) values of 0.7572% for the training phase and 0.8338% for the testing phase, accompanied by a noticeable reduction in mean absolute error (MAE). Furthermore, the model demonstrates strong robustness under noisy conditions, consistently outperforming the conventional FCM-RRBF approach. These findings highlight the relevance of combining adaptive fuzzy partitioning with global optimization techniques to achieve reliable and accurate SOC estimation in real-world battery applications.**

***Index Terms*—SOC Estimation, RRBF Networks, Fuzzy C-Means Clustering, Particle Swarm Optimization**

I. Introduction

The growing reliance on fossil fuels in transportation and communication systems—two major contributors to global energy demand and greenhouse gas emissions—has intensified the need for cleaner and more sustainable mobility solutions. In this regard, electric vehicles (EVs) have gained significant attention as a viable alternative. Among the various energy storage options, nickel–metal hydride (Ni–MH) batteries continue to be of interest due to their high safety level, long service life, and relatively low environmental impact.

Despite these advantages, accurate state of charge (SOC) estimation remains a critical challenge for battery management systems. Battery behavior is inherently nonlinear and strongly affected by factors such as temperature variations, aging, and operating conditions, which limits the effectiveness of conventional estimation methods. Techniques such as Coulomb counting and open-circuit voltage are widely adopted but are prone to cumulative errors and noise sensitivity, reducing their reliability in practical applications.

To enhance estimation accuracy, model-based strategies integrating equivalent circuit or electrochemical models with advanced observers have been developed. Although these approaches can improve performance, they require precise model identification and remain sensitive to parameter uncertainties. In parallel, data-driven methods have shown strong capabilities in capturing complex nonlinear dynamics directly from data, leveraging machine learning models such as deep neural networks, LSTM architectures, and support vector regression. However, their robustness often degrades under noisy environments and varying operating conditions.

More recently, hybrid intelligent approaches have emerged as an effective alternative by combining clustering, optimization, and learning mechanisms. In this context, recurrent radial basis function (RRBF) networks provide efficient nonlinear approximation with fast convergence, though their performance is highly dependent on proper initialization and parameter tuning. To mitigate these limitations, fuzzy c-means (FCM) clustering can be employed to enable adaptive data partitioning, while particle swarm optimization (PSO) offers an effective strategy for global parameter optimization and improved convergence.

Building on these concepts, this work proposes a hybrid PSO–AFCM–RRBF framework for accurate and robust SOC estimation. By integrating adaptive fuzzy clustering with global optimization, the proposed approach achieves a favorable balance between precision and robustness, particularly in noisy conditions, and demonstrates improved performance compared to conventional FCM–RRBF methods.

The remainder of this paper is structured as follows. Section II introduces the theoretical foundations of RRBF networks and related techniques. Section III details the proposed methodology. Section IV presents the experimental validation and comparative analysis. Finally, Section V concludes the study and outlines future research perspectives.

II. Recurrent Radial Basis Function Network (RRBF)

Recurrent Radial Basis Function (RRBF) networks constitute an efficient neural modeling framework that integrates recurrent dynamics with radial basis function approximation, enabling the representation of complex and time-varying nonlinear systems.

The architecture is composed of three main layers: an input layer incorporating local recurrent connections with sigmoid activation functions to capture temporal dependencies, a hidden layer formed by radial basis neurons for nonlinear transformation, and a linear output layer responsible for generating the final prediction. This configuration allows the model to jointly exploit temporal correlations and nonlinear relationships, as illustrated in Fig. 1.

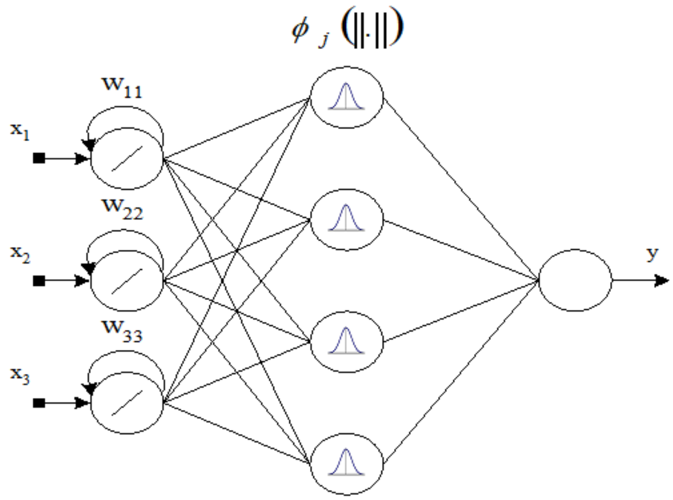

Fig. 1. Recurrent radial basis function network architecture.

The learning process is typically carried out in three steps: (i) adjustment of the recurrent input layer to encode temporal information, (ii) determination of the centers and spreads of the RBF neurons to ensure effective nonlinear representation, and (iii) estimation of the output weights using linear regression, ensuring computational efficiency.

The recurrent connections introduce a memory effect, while the sigmoid activation function

$$f(x) = \frac{1 - \exp(-kx)}{1 + \exp(-kx)} \tag{1}$$

provides a smooth nonlinear mapping, where $k > 0$ controls the sensitivity of the neuron.

At each discrete time instant t, the activation of the i-th recurrent neuron is defined by combining the current input $I_i(t)$ with its previous state $x_i(t-1)$ through a feedback weight w_{ii}:

$$a_i(t) = w_{ii}x_i(t-1) + I_i(t) \tag{2}$$

The corresponding neuron output is then obtained as:

$$X_i(t) = f(a_i(t)) \tag{3}$$

Assuming initial conditions $I_i(t_0) = 0$ and $X_i(t_0) = 1$, the memory depth can be controlled by the product kw_{ii}, which regulates the influence of past states. Under these assumptions, the neuron output can be expressed as:

$$X(t) = \frac{1 - \exp(-kw_{ii}x(t-1))}{1 + \exp(-kw_{ii}x(t-1))} \tag{4}$$

In the hidden layer, radial basis function neurons generate localized responses, allowing accurate approximation of nonlinear mappings. Each neuron is defined by a center v_j and a spread parameter σ_j, with activation given by:

$$\varphi_j = \exp\left(-\frac{1}{2\sigma_j^2}\|X_i - v_j\|^2\right) \tag{5}$$

The centers are commonly estimated as the mean of clustered samples:

$$v_j = \frac{1}{M}\sum_{i=1}^{M} X_i \tag{6}$$

while the spreads are derived from the sample dispersion:

$$\sigma_j^2 = \frac{1}{M-1}\sum_{i=1}^{M}(X_i - v_j)(X_i - v_j)^T \tag{7}$$

The final network output is computed as a linear combination of the hidden-layer activations:

$$\widehat{Y}_i = \sum_{j=1}^{n} w_j\, \varphi_j(X_i) \tag{8}$$

where w_j denotes the output weights and $\varphi_j(X_i)$ the response of the j-th radial basis neuron. This formulation enables the RRBF network to effectively combine temporal memory with localized nonlinear approximation, making it well suited for modeling dynamic nonlinear systems.

To further enhance robustness and generalization, appropriate initialization of RBF centers is essential. Clustering techniques such as K-means and Fuzzy C-Means (FCM) provide structured and representative prototypes, improving convergence behavior and predictive performance.

III. Fuzzy C-Means Clustering for RRBF

Fuzzy C-Means (FCM) is a soft partitioning method particularly suited to datasets characterized by overlapping structures or ambiguous boundaries. In contrast to hard clustering techniques, FCM assigns each data sample a degree of membership to multiple clusters, allowing a more flexible representation of data distribution. The algorithm iteratively updates cluster centers by minimizing a weighted objective function based on the similarity between samples and cluster prototypes:

$$J_{\text{FCM}} = \sum_{j=1}^{n}\sum_{i=1}^{M}(\mu_{ji})^m d_{ji}^2 \tag{9}$$

subject to

$$\sum_{j=1}^{n}\mu_{ji} = 1, \quad d_{ji}^2 = (x_i - v_j)^T(x_i - v_j), \tag{10}$$

where μ_{ji} denotes the membership degree of sample x_i in cluster j, $m > 1$ is the fuzzification parameter, and d_{ji} represents the Euclidean distance. Minimization of J_{FCM} leads to the following update rules:

$$\mu_{ji} = \frac{1}{\sum_{k=1}^{n} \left(\frac{d_{ji}}{d_{ki}}\right)^{\frac{2}{m-1}}} \tag{11}$$

$$v_j = \frac{\sum_{i=1}^{M} (\mu_{ji})^m x_i}{\sum_{i=1}^{M} (\mu_{ji})^m} \tag{12}$$

IV. Adaptive Fuzzy C-Means Clustering for RRBF

From Eq. (9), the standard FCM objective function can be rewritten as:

$$J_{FCM} = \sum_{i=1}^{n} \sum_{j=1}^{c} \mu_{ij}^m d_{ij}^2 \tag{13}$$

with $d_{ij}^2 = \|X_i - v_j\|^2 = (X_i - v_j)^T (X_i - v_j)$ representing the squared Euclidean distance.

This classical formulation assumes that all samples contribute equally during clustering, which may be inappropriate in the presence of noise, outliers, or nonstationary behaviors. To address this limitation, Adaptive Fuzzy C-Means (AFCM) introduces a weighting mechanism that adjusts the influence of each data point according to its relevance. This strategy improves robustness, enhances cluster separation in ambiguous regions, and better captures evolving data distributions. Furthermore, combining AFCM with optimization techniques such as Particle Swarm Optimization (PSO) enables improved parameter tuning and convergence stability.

The AFCM objective function is defined as:

$$J_{AFCM}(X, U, V) = \sum_{i=1}^{n} \beta_i^p \sum_{j=1}^{c} \mu_{ij}^m d_{ij}^2 \tag{14}$$

subject to:

$$\begin{cases} \sum_{j=1}^{c} \mu_{ij} = 1, \\ \prod_{i=1}^{n} \beta_i = 1, \end{cases} \tag{15}$$

where $\mu_{ij} \in [0, 1]$ denotes the membership degree of X_i to cluster j, and $\beta_i > 0$ is an adaptive weight controlling the contribution of each sample. The parameter p regulates the distribution of these weights, and the vector $\boldsymbol{\beta}$ reflects the global influence of all data points.

By minimizing J_{AFCM} using the Lagrange multiplier approach, the update expressions are obtained as:

$$\beta_i = \left[\frac{\left(\prod_{i=1}^{n} \sum_{j=1}^{c} \mu_{ij}^m d_{ij}^2 \right)^{\frac{1}{n}}}{\sum_{j=1}^{c} \mu_{ij}^m d_{ij}^2} \right]^{\frac{1}{p}} \tag{16}$$

$$v_j = \frac{\sum_{i=1}^{n} \beta_i \mu_{ij}^m X_i}{\sum_{i=1}^{n} \beta_i \mu_{ij}^m} \tag{17}$$

To ensure numerical stability, especially when distances $\|X_i - v_j\|^2$ are very close, a small tolerance ε (e.g., 10^{-10}) is introduced. In such cases, membership values are evenly distributed across clusters, avoiding numerical instability and ensuring consistent convergence.

Incorporating AFCM into the RRBF framework improves the quality of hidden-layer prototypes by limiting the influence of noisy or irrelevant samples. This leads to more representative cluster centers, improved coverage of the input space, and better localization of radial basis functions. Consequently, the overall model achieves higher approximation accuracy, reduced redundancy, and enhanced generalization performance when dealing with complex nonlinear dynamic systems.

V. Parameter Optimization of RRBF Networks Using Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population-based stochastic algorithm inspired by the collective behavior of social organisms. Each particle i represents a candidate solution with position $\mathbf{X}_{it}$ and velocity $\mathbf{V}_{it}$, updated using its personal best $\mathbf{P}_i$ and the global best $\mathbf{P}_g$:

$$\mathbf{V}_{i(t+1)} = w\mathbf{V}_{it} + c_1 r_1 (\mathbf{P}_i - \mathbf{X}_{it}) + c_2 r_2 (\mathbf{P}_g - \mathbf{X}_{it}) \tag{18}$$

$$\mathbf{X}_{i(t+1)} = \mathbf{X}_{it} + \mathbf{V}_{i(t+1)} \tag{19}$$

Here, w controls exploration versus exploitation, c_1 and c_2 are cognitive and social coefficients, and $r_1, r_2 \sim U[0, 1]$ introduce stochastic diversity.

For RRBF networks, PSO optimizes the centers and spreads of radial basis functions and other key parameters by minimizing a fitness function based on prediction error. Iteratively, particles refine their positions, guiding the swarm toward optimal solutions while reducing sensitivity to initialization and avoiding local minima.

VI. Analytical Learning of RRBF Networks Using Multiple Linear Regression

In this study, the output layer of the Recurrent Radial Basis Function (RRBF) network is trained using an analytical approach based on Multiple Linear Regression (MLR). This strategy enhances prediction accuracy while significantly reducing training time and improving robustness in the presence of noise and nonlinear dynamics. In contrast to iterative gradient-based methods, MLR offers a closed-form solution, ensuring numerical stability and avoiding issues such as slow convergence or divergence [1], [2]. The overall learning scheme combines three key elements: recurrent processing at the input layer, nonlinear mapping through radial basis functions, and analytical estimation of output weights.

The predicted output $\hat{Y}$ is defined as:

$$\hat{Y} = \phi \hat{W}, \quad \phi = \{\varphi_1, \varphi_2, \ldots, \varphi_n\} \tag{20}$$

The objective is to minimize the quadratic prediction error:

$$E = Y - \hat{Y} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_M \end{bmatrix}, \quad \min \sum_{i=1}^{M} e_i^2 = EE^T = (Y - \phi\hat{W})(Y - \phi\hat{W})^T \tag{21}$$

By differentiating with respect to $\hat{W}$ and equating to zero, the optimal solution is obtained as:

$$\frac{\partial(EE^T)}{\partial\hat{W}} = -2\phi^T Y + 2\phi^T \phi \hat{W} \tag{22}$$

$$\hat{W} = (\phi^\top \phi + \lambda I)^{-1} \phi^\top Y \tag{23}$$

where λI is an optional regularization term that improves numerical conditioning and limits overfitting.

This closed-form formulation enables fast and reliable estimation of the RRBF output weights, making it well suited for real-time state of charge (SOC) estimation in dynamic battery systems.

A. Dataset Characterization

The experimental dataset was obtained from tests conducted on commercial 18650 lithium-ion cells (NMC chemistry, 3 Ah, LG Chem). Cycling experiments were carried out using Arbin LBT21084 battery cyclers in combination with SPX Tenney T10C-1.5 thermal chambers, allowing precise temperature control over a wide range from -73 °C to 200 °C. The surface temperature of each cell was continuously measured using type-K thermocouples.

Before testing, all cells were allowed to reach thermal equilibrium for 24 h. The experimental protocol included different operating conditions, such as partial SOC ranges (20–80%, 40–60%) and full charge–discharge cycles (0–100%). Charging and discharging were performed under constant-current (CC) and constant-current–constant-voltage (CC–CV) modes. In addition, periodic reference tests at 0.5C, consisting of three full cycles, were conducted to recalibrate the SOC and monitor capacity degradation.

Key cell characteristics are reported in Table I. ICP-OES analysis revealed a nickel-rich cathode composition ($Ni_{0.84}Mn_{0.06}Co_{0.10}$), typical of NMC811 materials. A systematic design-of-experiments strategy was adopted to vary temperature, SOC range, and current rate. For each operating condition, at least two cells were tested to ensure reproducibility. The experiments were stopped when the capacity variation between cells exceeded 5%. Safety measures included automatic shutdown in the event of voltage deviations beyond ± 0.05 V or abnormal temperature increases.

The resulting dataset exhibits degradation patterns consistent with those reported in the literature, where capacity loss is primarily influenced by temperature, depth of discharge, and current rate [3]–[5]. After an initial stabilization phase related to solid electrolyte interphase (SEI) formation, degradation tends to follow an approximately linear trend driven by electrochemical and thermal effects.

To evaluate the robustness of the proposed model, two scenarios were considered: (i) ideal measurements without noise, and (ii) measurements affected by additive Gaussian noise with a signal-to-noise ratio (SNR) of 30 dB.

TABLE I
KEY SPECIFICATIONS OF THE 18650 LI-ION NMC CELLS

Property	Value / Description
Battery type	Lithium-ion 18650 (NMC)
Cathode material	NMC ($Ni_{0.84}Mn_{0.06}Co_{0.10}$)
Nominal capacity (Ah)	3
Nominal voltage (V)	3.6
Voltage range (V)	2.0 – 4.2
Maximum discharge current (A)	20
Operating temperature (°C)	-5 to 50

VII. PROPOSED APPROACH FOR SOC PREDICTION OF LI-ION EV BATTERIES

The proposed framework relies on a hybrid intelligent architecture that integrates adaptive fuzzy clustering, swarm-based optimization, and nonlinear learning to achieve accurate and robust SOC estimation.

First, adaptive fuzzy c-means (AFCM) clustering is employed to model the nonlinear electrochemical behavior of the battery through soft partitioning. To improve clustering quality and reduce sensitivity to initialization, particle swarm optimization (PSO) is used to refine cluster centers and membership degrees, enhancing convergence and stability.

The optimized centers are then used to construct radial basis function (RBF) kernels in the hidden layer. The output layer is formulated as a linear model, whose parameters are estimated via multiple linear regression (MLR), ensuring low computational complexity and reliable parameter estimation.

Overall, the integration of AFCM, PSO, and RRBF with an MLR-based output layer results in a compact and robust framework, offering high prediction accuracy and strong generalization capability. The complete learning procedure is summarized in Algorithm 1. The main parameters and hyperparameters of the proposed PSO–AFCM–RRBF framework are summarized in table II to ensure reproducibility and clarity of the experimental setup.

Algorithm 1: Proposed PSO–AFCM–RRBF Algorithm for SOC Estimation

Input: Dataset $\mathbf{P}$, target $\mathbf{f}$, PSO–AFCM parameters, epochs E
Output: Estimated SOC $\hat{\mathbf{f}}$, trained parameters $\{w, b, w_1\}$

Step 1: Data Preprocessing
Load dataset and extract features $\mathbf{P}$ and target $\mathbf{f}$.
Convert energy to capacity: $f_{Ah} = \frac{f_{Wh}}{V}$.
Split data into training and testing sets.

Step 2: Clustering using PSO-AFCM
Initialize swarm particles and AFCM parameters.
for *each iteration* **do**
    Update particle velocities and positions.
    Compute clustering objective function J.
    Update cluster centers $\mathbf{C}$ and fuzziness coefficient m.
Obtain optimal centers $\mathbf{C}$ and membership matrix $\mathbf{U}$.

Step 3: Initialize RRBF Parameters
Initialize weights w, bias b, and input weights w_1 randomly.

Step 4: Training Phase
for *each run* **do**
    for *each epoch* $k = 1 \to E$ **do**
        Initialize sliding window vector $\mathbf{U}_{vec}$.
        for *each sample* i **do**
            compute the output of input layer using equation (4) Compute RBF activation using eq (5):
            Compute network output using eq (8)
        Update parameters (MLR) using eq (23):

Step 5: Testing Phase
for *each test sample* **do**
    Apply sliding window.
    Compute hidden activations ϕ_{test}.
    Estimate output $\hat{f}_{test}$.

Step 6: Performance Evaluation
Compute RMSE and NRMSE for training and testing:
return Estimated SOC and performance metrics.

A. Performance Evaluation of the Proposed Framework

This subsection presents a comprehensive evaluation of the proposed PSO–AFCM–RRBF framework for SOC estimation. Multiple hybrid configurations are tested to quantify the contribution of each component in terms of accuracy, robustness, and computational efficiency under both noise-free and noisy conditions.

For benchmarking, a baseline RRBF network is implemented with prototypes initialized using FCM clustering. The performance of the proposed PSO–AFCM–RRBF framework is illustrated in Figs. 2 and 3 under clean and noisy scenarios, respectively. These results highlight the improvements achieved through PSO-based cluster optimization and adaptive fuzzy activations, demonstrating enhanced estimation accuracy and robustness compared to conventional FCM-RRBF approaches.

TABLE II
PARAMETERS OF THE PROPOSED AFCM–PSO–RRBF FRAMEWORK

Parameter	Symbol	Value
Number of clusters	C	5
Fuzzification coefficient	m	[1.5–2.5] (optimized)
Maximum iterations	$T_{\max}$	500
Inertia weight	w	[0.4–0.9]
Cognitive coefficient	c_1	1.8
Social coefficient	c_2	1.8
Velocity limit	$v_{\max}$	0.2
Number of hidden neurons	N_h	5
Epochs	E	50

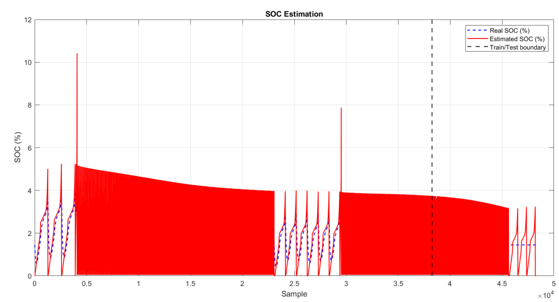

Fig. 2. SOC Estimation Using the Proposed PSO–AFCM–RRBF Framework

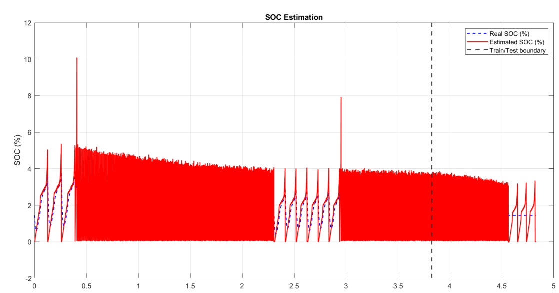

Fig. 3. SOC Estimation under Noise Using the Proposed Hybrid Model

TABLE III
COMPARATIVE PERFORMANCE OF DIFFERENT INTELLIGENT FRAMEWORKS FOR SOC ESTIMATION.

Phase	Metric	FCM-RRBF	Proposed Method
SOC Training	RMSE (%)	0.8543	**0.7572**
	MAE (%)	0.7520	**0.6208**
SOC Testing	RMSE (%)	0.9052	**0.8338**
	MAE (%)	0.7551	**0.6448**

TABLE IV
COMPARATIVE SOC ESTIMATION PERFORMANCE OF HYBRID MODELS UNDER NOISE

Phase	Metric	FCM-RRBF	Proposed Method
SOC Training	RMSE (%)	0.8597	**0.7577**
	MAE (%)	0.7230	**0.6210**
SOC Testing	RMSE (%)	0.9438	**0.8342**
	MAE (%)	0.7548	**0.6448**

VIII. DISCUSSION

The obtained results clearly demonstrate the effectiveness of the proposed PSO-AFCM-RRBF framework for SOC estimation under both nominal and noisy operating conditions. The observed improvement in accuracy, as reflected by lower RMSE and MAE values, can be attributed to the complementary strengths of each component in the hybrid architecture.

Specifically, the AFCM clustering provides a flexible representation of the nonlinear battery dynamics by preserving uncertainty through soft partitioning. The integration of PSO further enhances this process by optimizing cluster centers and membership distributions, thereby reducing sensitivity to initialization and improving convergence toward globally relevant solutions. This directly impacts the quality of the feature space used by the RRBF network.

Moreover, the RRBF model benefits from the optimized clustering structure, leading to improved generalization capability and reduced approximation error. Compared to the baseline FCM–RRBF approach, the proposed framework exhibits a more stable behavior in the presence of measurement noise. This robustness is particularly important for real-world battery management systems, where sensor noise and operating variability are unavoidable.

Another key observation is the balanced trade-off achieved between model complexity and performance. While more sophisticated deep learning models may offer high accuracy, they often require extensive training data and computational resources. In contrast, the proposed method maintains a compact structure with efficient training, making it suitable for real-time embedded applications.

Overall, the results highlight that the joint optimization of clustering and learning stages is essential for improving SOC estimation reliability. The proposed framework not only enhances prediction accuracy but also ensures robustness and computational efficiency, which are critical requirements for practical electric vehicle applications.

IX. CONCLUSION

A hybrid PSO–AFCM–RRBF framework for lithium-ion battery SOC estimation has been proposed. By combining adaptive fuzzy clustering, swarm-based optimization, and a nonlinear learning model, the approach improves both estimation accuracy and robustness.

Experimental results show that the proposed method outperforms the FCM–RRBF baseline in terms of RMSE and MAE for both training and testing phases. In addition, it maintains stable performance under noisy conditions, demonstrating enhanced robustness.

These findings confirm the effectiveness of integrating fuzzy clustering with global optimization for reliable SOC estimation. Future work will address real-time implementation and the extension to joint SOC/SOH estimation.

REFERENCES

[1] Zuzana Majdisova and Vaclav Skala. Radial basis function approximations: comparison and applications. *Applied Mathematical Modelling*, 51:728–743, 2017.

[2] Ersheng Pan, Dong Peng, Wangcheng Long, Yawei Xue, Lang Zhao, and Jinchao Li. Provincial grid investment scale forecasting based on mlr and rbf neural network. *Mathematical Problems in Engineering*, 2019(1):3197595, 2019.

[3] Yuliya Preger, Heather M Barkholtz, Armando Fresquez, Daniel L Campbell, Benjamin W Juba, Jessica Romàn-Kustas, Summer R Ferreira, and Babu Chalamala. Degradation of commercial lithium-ion cells as a function of chemistry and cycling conditions. *Journal of The Electrochemical Society*, 167(12):120532, 2020.

[4] Andrew Weng, Peyman Mohtat, Peter M Attia, Valentin Sulzer, Suhak Lee, Greg Less, and Anna Stefanopoulou. Predicting the impact of formation protocols on battery lifetime immediately after manufacturing. *Joule*, 5(11):2971–2992, 2021.

[5] Cunzhi Zhao, Xingpeng Li, and Yan Yao. Quality analysis of battery degradation models with real battery aging experiment data. In *2023 IEEE Texas Power and Energy Conference (TPEC)*, pages 1–5. IEEE, 2023.