

ORIGINAL RESEARCH

Open access

# An Intelligent Framework for Electric Vehicle Charging Duration Estimation Based on Deep Feature Learning and Clustering

Lucia Romano<sup>1\*</sup>, Marco Bellini<sup>2</sup>, Paolo Greco<sup>1</sup>

## Abstract

Reliable forecasting of electric vehicle (EV) charging time is vital for improving driver convenience and enabling smart, real-time charge scheduling. The present work introduces a comprehensive dynamic prediction system for EV charging duration that integrates four complementary modules: data preparation, pattern identification, baseline forecasting, and adaptive error adjustment. Raw measurements from the Battery Management System (BMS) are first cleaned and scaled to correct for missing or erroneous entries. A specially designed convolutional autoencoder tailored for EVs (EV-CAE) then extracts features across multiple time scales, after which K-Means clustering groups the sessions into characteristic charging profiles. Drawing on these profile assignments, the baseline forecasting stage computes an initial estimate of remaining charging time using past session records and the assigned cluster label. For improved resilience to changing circumstances, an adaptive correction layer is incorporated; it blends linear regression, logarithmic fitting, ratio-based scaling, and neural-network-driven refinement to update the forecast during the session continuously. Validation on authentic EV field datasets reveals that the overall architecture delivers a clear gain in forecast quality. In particular, activating the adaptive correction layer increases the coefficient of determination ( $R^2$ ) from 0.948 to 0.960. It maintains consistent accuracy even when charging curves vary unexpectedly or when operating in cold-weather conditions. These outcomes highlight the framework's readiness for deployment and its engineering relevance in next-generation intelligent EV charging infrastructure.

**Keywords** Electric vehicles, Charging duration prediction, Charging pattern recognition, Dynamic bias correction, EV-CAE

\*Correspondence:

Lucia Romano

lucia.romano@uniroma1.it

<sup>1</sup> Department of Electrical Power Systems, Faculty of Engineering, Sapienza University of Rome, Rome, Italy

<sup>2</sup> Department of Energy Informatics, Faculty of Engineering, Polytechnic University of Milan, Milan, Italy

## Introduction

### Research background

The accelerating worldwide shift toward electric vehicles (EVs) has placed charging infrastructure and user experience at the center of current research efforts. The International Energy Agency's World Energy Outlook 2024 indicates that EVs already account for roughly 20% of all newly registered passenger vehicles globally. Under prevailing policy trajectories, that figure is projected to approach 50% by 2030. This pronounced expansion over the next 10 years is poised to fundamentally alter the automotive market landscape and lead to a major reduction in petroleum demand for surface transportation [1]. Within this evolving environment, delivering a seamless, predictable charging process has emerged as

a decisive factor in shaping consumer willingness to adopt EVs, with the length of charging sessions standing out as a primary source of user dissatisfaction.

A substantial body of earlier work has centered on the charging-station side, building predictive models that draw on live data streams captured at the station [2, 3]. Although these investigations have yielded meaningful progress, the highly uneven rollout of advanced, data-rich charging facilities—predominantly located in densely populated metropolitan zones and largely absent from outlying districts—continues to limit widespread applicability. In sparsely populated or rural territories, insufficient data availability severely hampers the practical rollout of sophisticated prediction techniques [4]. This geographic imbalance emphasizes the value of shifting the analytical focus toward onboard vehicle data. A vehicle-side prediction strategy circumvents many infrastructure-related constraints and carries important real-world utility.

Establishing a robust, vehicle-centric method capable of accurately estimating charging duration in representative real-life situations—including extended highway trips, overnight residential slow charging, and off-grid or underserved locations where station-derived information is missing or untrustworthy—remains an open and pressing problem. To address this need, the current study presents a high-performance architecture for dynamic charging-duration prediction. Beyond generating estimates that account for different starting state-of-charge (SOC) values, the system continuously tracks and adapts to the evolving charging curve in real time. Relative to conventional approaches dependent on station-side measurements, the suggested framework offers several distinct strengths: (i) it exploits data directly from the vehicle's own sensors and controllers, thereby mirroring the true physical and environmental state of the EV more faithfully; (ii) it consistently achieves elevated levels of predictive precision; (iii) it incorporates live adaptation capabilities that automatically respond to deviations observed during the charging event; and (iv) it maintains strong stability against irregular or anomalous behavior in the charging profile, thereby enhancing end-to-end trustworthiness. Empirical testing confirms that the architecture excels at both categorizing charging patterns and forecasting duration, attaining a coefficient of determination ( $R^2$ , as defined herein) of 0.96. This performance substantially narrows prediction uncertainty and furnishes dependable algorithmic foundations for real-time charge supervision and time-to-complete estimation in intelligent EV ecosystems.

## Research status and challenges

The core objective of this work is to enhance both the precision and the flexibility of electric vehicle (EV) charging duration forecasting. To this end, the study proposes a unified framework that integrates several interconnected techniques, with charging pattern recognition serving primarily as an auxiliary process rather than the central aim. The principal difficulties and underlying technical groundwork tackled in this research are outlined as follows:

### (1) Charging Pattern Recognition as an Enabling Component:

Charging behavior in EVs shows considerable variety, with charging profiles differing markedly across usage contexts. Correctly identifying the prevailing charging pattern enables the prediction model to better align with the specifics of the ongoing session, thereby boosting overall accuracy. Although pattern recognition itself is not the ultimate target here, it serves a vital supporting function by mitigating data diversity and laying the groundwork for more effective subsequent forecasting. Current research shows no universal agreement on classification standards, which generally fall into two broad approaches:

#### a. Classification Based on Physical Characteristics:

Initial investigations divided the charging process into distinct stages, most commonly the constant current (CC) and constant voltage (CV) phases [5]; these divisions served as the basis for many control and optimization strategies. Yet the emergence of more sophisticated charging protocols—including multistage CC-CV, temperature-dependent adjustments, and dynamically varying current profiles—has rendered such basic categorizations inadequate for representing the wide range of behaviors seen across different vehicles and charging equipment [6]. Efforts to standardize charging categories

by power tier (e.g., Level 1 through Level 4 [7]) can be useful for high-level infrastructure planning, but they offer coarse granularity and inconsistent adoption worldwide [8]. Furthermore, these static, physics-driven labels are inherently unable to respond to live shifts in battery health, temperature, or other environmental conditions. For instance, refinements to the CC-CV approach suggested by Koli et al. [9] and Sabarimuthu et al. [10] offered incremental advances. Yet, their applicability under diverse and realistic operating scenarios remained largely unvalidated.

#### b. Classification Based on Charging Duration:

Certain studies segment charging events by duration—typically into slow, fast, and ultra-fast categories—as examined by Suarez and Martinez [8] and Madaram et al. [11]. Although straightforward and user-friendly, these groupings differ across publications and overlook critical vehicle-specific attributes and external influences. These shortcomings highlight the advantages of adopting data-driven clustering techniques that can adapt to variations in vehicle models, charger specifications, and behavioral patterns. Within the present framework, clustering and pattern detection are applied exclusively to create more uniform subsets of the training data, which, in turn, supports greater accuracy in the main duration prediction task.

#### (2) Charging Duration Prediction:

Traditional approaches depended on piecewise or physics-informed models of the charging curve [12–14]. These methods yielded satisfactory results in controlled settings. Still, they proved fragile when confronted with the irregularities and noise inherent in field-collected data, owing to their dependence on rigid assumptions and high-quality inputs. The incorporation of machine learning and deep learning has brought substantial gains in robustness. Earlier machine learning techniques, such as Support Vector Machines (SVMs) and Random Forests (RFs) [15], delivered promising initial results. At the same time, subsequent developments leveraging Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks [16, 17] have established leading performance in time-series forecasting. More advanced configurations, including transformer models [3] and TCN-LSTM combinations [18], have further raised the performance ceiling. However, their high computational demands often make them impractical for onboard, real-time execution in vehicles. Kreft et al. [19], drawing on extensive residential charging datasets, emphasized the wide range of user habits and the shortcomings of uniform, centralized prediction schemes—reinforcing the value of pattern-sensitive approaches to improve forecast quality. Accurate duration estimation remains especially challenging in the presence of rapid fluctuations and external disturbances.

#### (3) Strategies for Handling External Interference Factors:

Factors such as variations in grid demand, changes in ambient temperature, and unpredictable user actions can significantly degrade prediction reliability. Existing countermeasures typically involve either post-facto adjustments (e.g., residual-based corrections [14]) or continuous online refinement, including techniques such as Kalman filtering [20] and distributed learning frameworks at the edge [21]. Nonetheless, achieving a satisfactory trade-off between instantaneous responsiveness and acceptable computational overhead remains unresolved.

Overall, real-world charging duration prediction faces persistent obstacles stemming from the lack of standardized pattern definitions, insufficient resilience to external perturbations, and difficulties with scalable, vehicle-side implementation. While charging pattern recognition is not the primary focus of this research, it is deliberately leveraged as a critical supporting tool to strengthen model generalization and minimize systematic forecasting errors. The introduced framework integrates automated pattern clustering, baseline static prediction, and real-time dynamic bias correction into a cohesive workflow, thereby overcoming the identified limitations and enabling accurate, responsive, and practically deployable charging duration estimation for EVs.

## Research objectives

To boost both the precision and real-world viability of electric vehicle (EV) charging duration forecasting, this investigation centers on two fundamental scientific questions. First, is it feasible to automatically categorize EV charging behaviors in the absence of pre-set category counts or uniform annotation standards? Second, can real-time interventions effectively counteract external disruptions to achieve precise, instantaneous charging duration estimates?

To tackle these inquiries, the study advances an integrated prediction architecture that merges pattern detection, baseline modeling, and adaptive refinement. The specific research goals include:

(1) Establishment of a dynamic EV charging duration prediction framework.

An all-encompassing prediction system is engineered, featuring a baseline modeling component and an adaptive refinement component. The baseline component delivers an initial duration estimate drawing on historical attributes (e.g., SOC, voltage, current), whereas the adaptive component applies ongoing adjustments informed by live sensor inputs. This architecture strengthens responsiveness and on-the-fly adaptability in variable environments, making it suitable for a broad array of charging contexts.

(2) Introduction of an unsupervised charging pattern recognition approach grounded in deep autoencoding and clustering.

A specialized convolutional autoencoder (EV-CAE) is developed to extract multi-scale temporal characteristics from charging sequences. K-Means clustering is then employed on these characteristics to enable unsupervised grouping of charging behaviors. Requiring neither prior domain expertise nor annotated examples, the technique exhibits robust generalization and flexibility, thereby producing more refined datasets to support the subsequent baseline prediction stage.

(3) Formulation of diverse bias correction approaches to strengthen forecasting resilience.

To counteract uncertainties like abrupt current variations during charging, a dedicated bias correction component is developed, encompassing four distinct methodologies: linear, logarithmic, proportional, and deep learning-driven. The component supports flexible switching between methodologies and intelligently selects the most suitable correction tactic based on observed residual trends in real time. This design elevates stability and versatility in intricate settings while facilitating implementation on vehicle-embedded edge computing platforms.

## Materials and Methods

This section begins with a high-level description of the proposed dynamic charging duration prediction framework for EVs, then provides an in-depth account of the functioning and sequence of operations for each constituent module.

### Framework overview

As depicted in **Figure 1**, the suggested framework comprises four primary modules: data preparation, charging pattern classification, static prediction, and dynamic bias correction. These modules operate synergistically in a closed-loop structure, facilitating the progression from historical data analysis to live estimation of remaining charging time. It should be emphasized that the input data originates from the Battery Management System (BMS) onboard EVs and is prone to issues such as gaps or irregular readings. Consequently, the dataset undergoes preprocessing steps—including format standardization, imputation of missing values, outlier mitigation, and scaling—to ensure reliable inputs for subsequent stages.

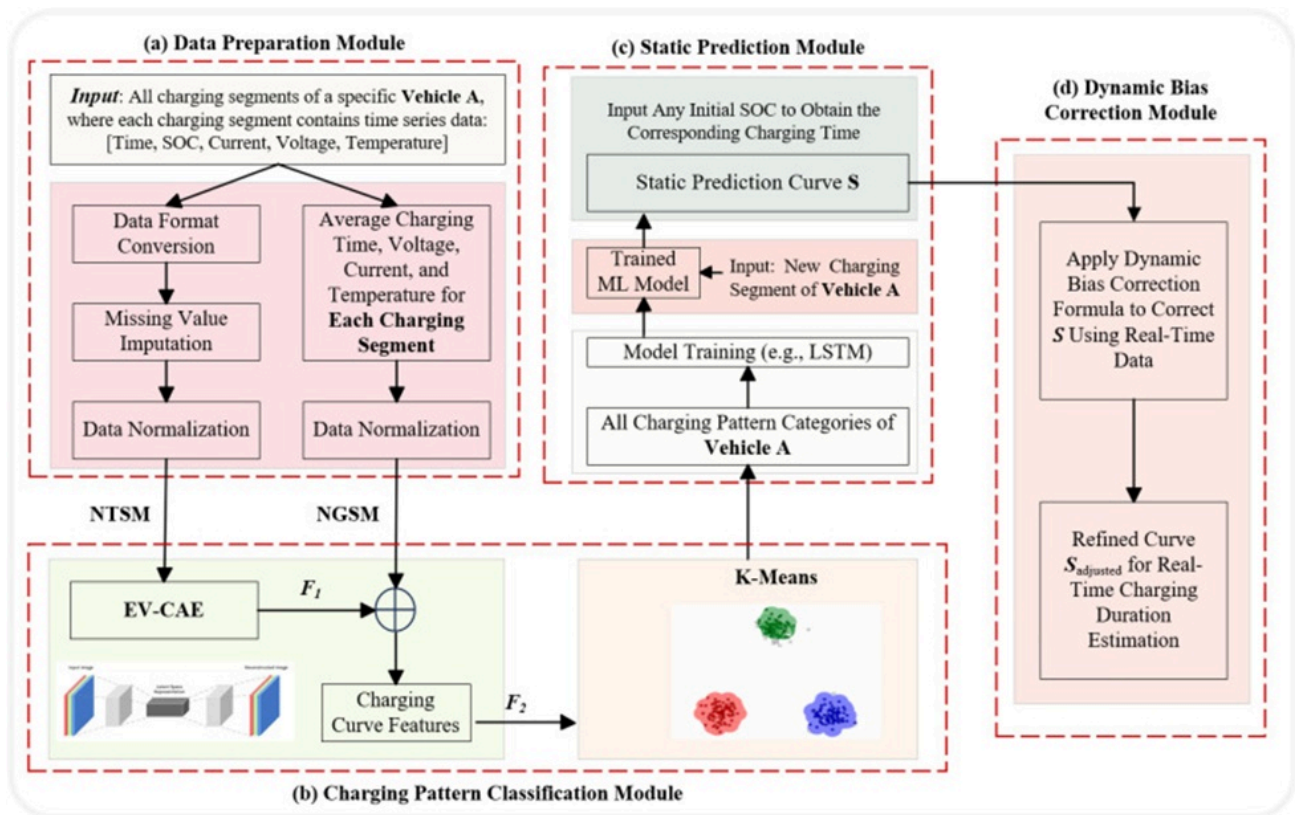


**Figure 1.** Dynamic charging duration prediction framework for EVs, comprising four modules: (a) Data Preparation Module, (b) Charging Pattern Classification Module, (c) Static Prediction Module, and (d) Dynamic Bias Correction Module.

The operational sequence of the prediction framework is exemplified using Vehicle A. Initially, every charging segment is uniformly resized to a consistent length. Missing data points are imputed using low-rank matrix completion techniques, and the entire set is normalized to yield time-series sequences along with comprehensive statistical summaries suitable for modeling. Next, the charging pattern classification module employs a convolutional architecture to derive features at multiple temporal resolutions, integrates them with global statistical summaries, and uses K-Means clustering to partition historical charging segments into representative categories.

During the static prediction phase, a sequence modeling architecture (e.g., LSTM) is trained on archived data along with their associated pattern assignments. Upon arrival of a fresh charging segment from Vehicle A, the system first assigns its pattern label and subsequently constructs an idealized charging trajectory  $S$  (the static prediction trajectory) derived from the mean properties of comparable historical segments within the same cluster. From any given starting SOC level, the complete charging duration can then be derived from  $S$ .

Nevertheless, the static prediction trajectory overlooks instantaneous variations, such as current instability. To overcome this shortcoming, a dynamic bias correction module is incorporated. Leveraging real-time observations from the immediately prior timestep, it applies adjustment equations to produce a revised forecast for the forthcoming timestep, yielding a continuously refined trajectory  $S_{\text{adjusted}}$ . Through ongoing application, this module progressively minimizes discrepancies between predicted and observed values, thereby aligning more closely with the actual advancement of the charging session.

## Data preparation module

The features selected for this investigation encompass time, current, voltage, temperature, and SOC. The dataset comprises a series of continuous charging segments, each containing sampled timestamps and corresponding values for current, voltage, temperature, and SOC. To clarify the processing steps, the description below employs “Car A” as a concrete illustration.

Suppose Car A possesses a total of  $n$  charging segments, collectively represented as  $CA = \{C_1, C_2, \dots, C_i, \dots, C_n\}$ . In this notation,  $CA$  denotes the complete collection of charging segments for Car A,  $C_i$  denotes the  $i$ -th charging segment, and  $n$  denotes the total number of such segments for Car A. Each charging segment is composed of a sequence of time-series observations. Considering segment  $C_i$  specifically, it can be formulated as  $C_i = \{(t_{ij}, I_{ij}, V_{ij}, T_{ij}, s_{ij}) \mid j = 1, 2, \dots, N_i\}$ , where  $t_{ij}$  is the sampling timestamp,  $I_{ij}$  is the recorded current,  $V_{ij}$  is the recorded voltage,  $T_{ij}$  is the recorded temperature, and  $s_{ij}$  is the recorded SOC value.  $N_i$  denotes the total number of sampling points present in segment  $C_i$ .

To better capture the essential patterns of SOC progression, each charging segment is subdivided into SOC increments. Again using  $C_i$  as the example, the boundary indices marking SOC transitions are denoted as  $j_1, j_2, \dots, j_{K_i+1}$ , with  $j_1 = 1$  and  $j_{K_i+1} = N_i + 1$ . Using these boundary points, the segment  $C_i$  is split into  $K_i$  sub-segments, each spanning the range from one SOC threshold to the next. For the  $k$ -th SOC interval (where  $k = 1, 2, \dots, K_i$ ), the statistical characteristics computed over that interval are defined as follows:

$$\Delta t_i^{(k)} = t_{i,j_{k+1}} - t_{i,j_k} \quad (1)$$

$$\bar{I}_i^{(k)} = \frac{1}{n_i^{(k)}} \sum_{j=j_k}^{j_{k+1}-1} I_{ij} \quad (2)$$

$$\bar{V}_i^{(k)} = \frac{1}{n_i^{(k)}} \sum_{j=j_k}^{j_{k+1}-1} V_{ij} \quad (3)$$

$$\bar{T}_i^{(k)} = \frac{1}{n_i^{(k)}} \sum_{j=j_k}^{j_{k+1}-1} T_{ij} \quad (4)$$

Where  $\Delta t_i^{(k)}$  signifies the length of the corresponding time span,  $\bar{I}_i^{(k)}$  represents the mean value of current,  $\bar{V}_i^{(k)}$  represents the mean value of voltage,  $\bar{T}_i^{(k)}$  represents the mean value of temperature, and  $n_i^{(k)} = j_{k+1} - j_k$  gives the quantity of sampling points located inside that particular SOC interval.

As a result, every charging segment  $C_i$  can be recast as an ordered collection of these computed statistical indicators, written as  $C'_i = \{\Delta t_i^{(k)}, \bar{I}_i^{(k)}, \bar{V}_i^{(k)}, \bar{T}_i^{(k)} \mid k = 1, 2, \dots, K_i\}$ . The resulting representation serves as the essential input for all subsequent stages and is designated as the raw transformed data. The aggregated set of raw-transformed data covering all charging segments from Car A is denoted  $C^*A = \{C'_1, C'_2, \dots, C'_i, \dots, C'_n\}$ , where  $n$  denotes the total number of charging segments associated with Car A, and  $C^*A$  denotes the entire ensemble of transformed segments.

Given that the number of intervals  $K_i$  differs from one transformed segment  $C'_i$  to another, standardization of segment length is required by consistently elongating each one (for example, extending them so that they uniformly cover the SOC progression from 0 up to 100). This procedure yields matrices having identical dimensions. Thus, the fully transformed dataset from all charging segments can be arranged into a single matrix  $R_{n \times d}$ , where  $n$  is the total number of charging segments and  $d$  is the selected uniform dimension.

In this research, a soft-thresholding approach based on low-rank matrix completion is utilized to impute the missing elements in  $M$ , to resolve the subsequent optimization formulation:

$$\min_X \|\Omega \odot (M - X)\|_F^2 + \lambda \|X\|_* \quad (5)$$

Where  $X$  denotes the completed (imputed) matrix,  $\Omega$  serves as a binary selection matrix (assigning 1 to positions where entries in  $M$  are observed and 0 to missing positions),  $\|\cdot\|_F$  indicates the Frobenius norm, and  $\|\cdot\|_*$  stands for the nuclear norm of the matrix (defined as the total of its singular values). The parameter  $\lambda$  functions as the regularization coefficient, controlling the balance between fidelity to the observed data and the low-rank constraint imposed on the solution. By pursuing this minimization, the approach successfully imputes missing entries via nuclear-norm reduction while exerting negligible influence on the subsequent clustering procedure.

Following completion, every feature column within the imputed matrix  $X$ —for instance, the columns corresponding to time interval, mean current, mean voltage, and mean temperature—is subjected to z-score standardization:

$$\mathcal{J}_i^{(a)} = \frac{x_i^{(a)} - \mu_a}{\sigma_a} \quad (6)$$

Where  $\mathcal{T}(a)_i$  indicates the normalized (standardized) value corresponding to the  $i$ -th charging segment for the  $a$ -th feature,  $x(a)_i$  is the unprocessed original value for that same entry, and  $\mu_a$  together with  $\sigma_a$  represent, respectively, the mean and the standard deviation computed across all values of the  $a$ -th feature.

In the final step, the fully standardized time-series feature matrix is referred to as the Normalized Time-Series Matrix (NTSM) and is expressed in the following form:

$$\text{NTSM} = \{\mathcal{J}_1, \mathcal{J}_2, \dots, \mathcal{J}_i, \dots, \mathcal{J}_n\} \quad (7)$$

Where NTSM stands for the complete collection of standardized charging segments about Car A, and  $\mathcal{T}_i$  refers to the standardized feature matrix associated with the  $i$ -th charging segment, formulated as:

$$\mathcal{J}_i = [\mathcal{J}_i^{(1)}, \mathcal{J}_i^{(2)}, \dots, \mathcal{J}_i^{(a)}, \dots] \quad (8)$$

Additionally, for each transformed charging segment  $C_i$ , a set of global-average features  $f_i$  can be derived, representing the segment's overall statistical summary. This is formulated as:

$$\bar{t}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \Delta t_i^{(k)} \quad (9)$$

$$\bar{I}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \bar{I}_i^{(k)} \quad (10)$$

$$\bar{V}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \bar{V}_i^{(k)} \quad (11)$$

$$\bar{T}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \bar{T}_i^{(k)} \quad (12)$$

Where  $K_i$  indicates the total count of SOC intervals present within the transformed charging segment  $C_i$ , while  $\Delta t(k)_i$ ,  $\bar{I}(k)_i$ ,  $\bar{V}(k)_i$ , and  $\bar{T}(k)_i$  correspond, respectively, to the duration of the time span, the mean current, the mean voltage, and the mean temperature recorded over the  $k$ -th SOC interval.

In this manner, the global average feature vector summarizing the entire charging segment  $C_i$  takes the form  $f_i = [\bar{t}_i, \bar{I}_i, \bar{V}_i, \bar{T}_i]$ . Accordingly, the complete collection of global statistical features across all charging segments of Car A is organized into what is termed the Global Statistical Matrix (GSM), expressed as  $GSM = \{f_i = [\bar{t}_i, \bar{I}_i, \bar{V}_i, \bar{T}_i] \mid i = 1, 2, \dots, n\}$ , with  $n$  denoting the overall number of charging segments for Car A. To place every feature on an equivalent scale, z-score normalization is subsequently performed on each feature column:

$$z_i^{(a)} = \frac{f_i^{(a)} - \mu'_a}{\sigma'_a} \quad (13)$$

Where  $z_{ai}$  denotes the standardized (normalized) value assigned to the  $i$ -th charging segment with respect to the  $a$ -th feature,  $f_i^{(a)}$  corresponds to the unnormalized original measurement for that entry, and  $\mu'_a$ , along with  $\sigma'_a$ , represent, respectively, the mean and the standard deviation calculated over all charging segments for the  $a$ -th feature.

Ultimately, these standardized global statistical features are collectively termed the Normalized Global Statistical Matrix (NGSM) and are expressed in the following manner:

$$NGSM = \{z_1, z_2, \dots, z_i, z_n\} \quad (14)$$

Where NGSM denotes the complete collection of standardized global features encompassing every charging segment from Car A, and  $z_i$  refers to the standardized feature vector that aggregates all global statistical measures specifically for the  $i$ -th charging segment, given by:

$$z_i = \left[ z_i^{(1)}, z_i^{(2)}, \dots, z_i^a, \dots \right] \quad (15)$$

## Charging pattern classification module

To enable reliable grouping of charging segment behaviors and lay a robust groundwork for the ensuing duration estimation procedures, the present work develops a specialized classification technique that combines a tailored convolutional autoencoder for EVs (EV-CAE) with subsequent clustering. The complete design of this component is outlined in [Figure 2](#).

**Figure 2(a)** illustrates the core network layout, which features three concurrent 1D convolutional branches employing filter widths of  $5 \times 1$ ,  $7 \times 1$ , and  $9 \times 1$ . These branches are specifically intended to pull out temporal patterns at varying resolutions from the Normalized Time-Series Matrix (NTSM). Once extracted, the multi-resolution feature sets pass through Batch Normalization to stabilize distributions, then undergo ReLU nonlinearity to strengthen expressive capacity. The diverse feature representations are subsequently merged along the channel dimension and processed by a dense (linear) transformation layer, yielding the compact high-level descriptor  $F_1$ .

In addition, the subfigures provide detailed breakdowns of key operations: **Figure 2(a-1)** to **Figure 2(a-3)** visualize how each convolutional kernel performs its sliding-window convolution across different receptive fields; **Figure 2(a-4)** clarifies the internal workings of Batch Normalization; **Figure 2(a-5)** graphs the piecewise-linear characteristics and derivative of the ReLU function; and **Figure 2(a-6)** diagrams the architecture and purpose of the Linear (fully connected) projection layer.

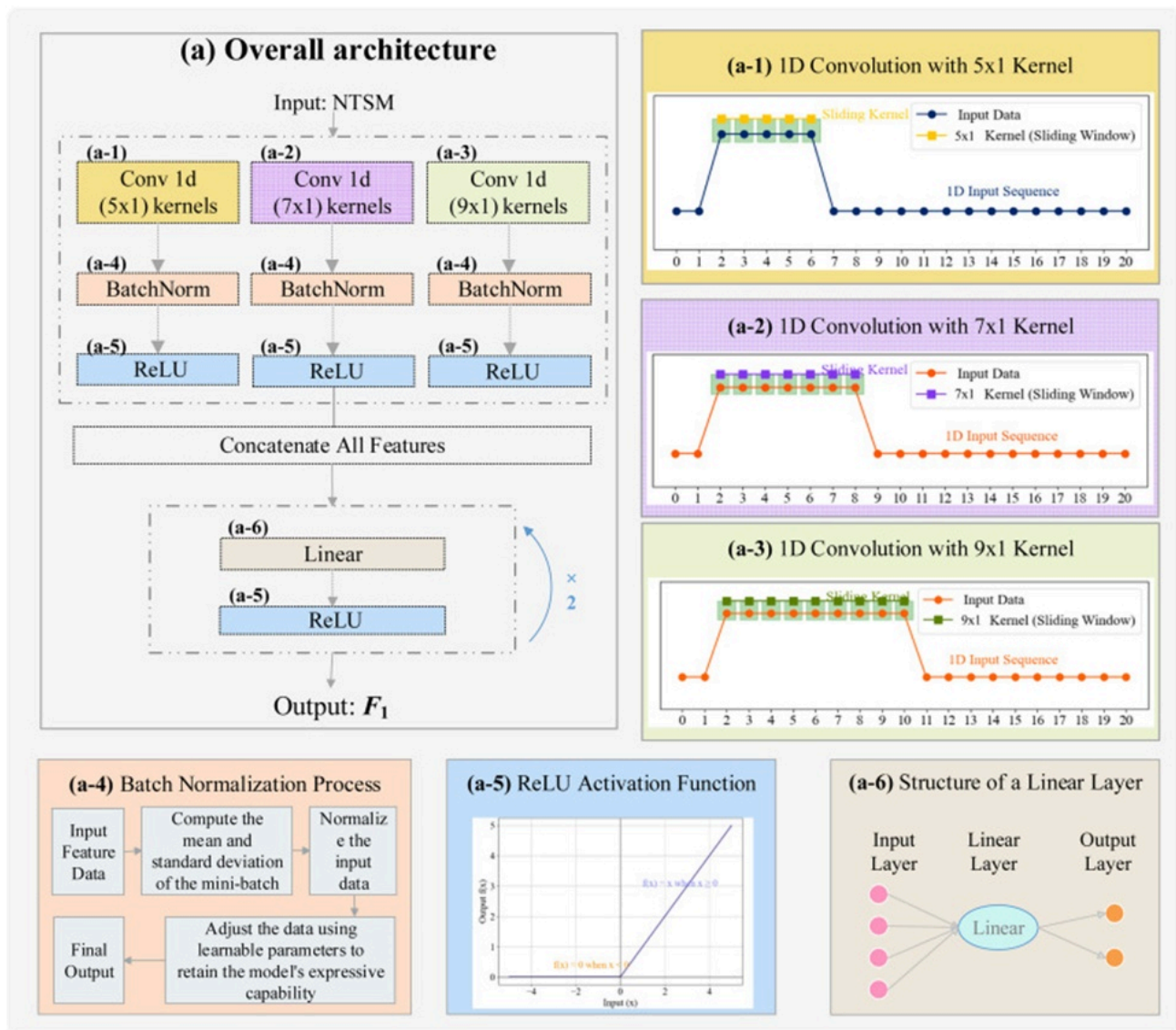


**Figure 2.** Architecture and principal computational steps of the EV-CAE: (a) complete model layout, (a-1)–(a-3) one-dimensional convolution operations using kernel sizes of  $5 \times 1$ ,  $7 \times 1$ , and  $9 \times 1$ , respectively, (a-4) Batch Normalization procedure, (a-5) ReLU nonlinearity, (a-6) schematic of the Linear (dense) layer.

Suppose the  $i$ -th charging segment belonging to Car A is denoted  $C_i$ . The following expression mathematically describes the convolutional computation applied to this segment during the feature extraction phase:

$$H_i = [h_i^{(5)}, h_i^{(7)}, h_i^{(9)}], h_i^{(r)} = \max\left(0, \frac{W_r * C_i + b_r - \mu_r''}{\sigma_r''}\right) \quad (16)$$

Where  $W_r$  stands for the convolutional filter with dimensions  $r \times 1$ , where  $r$  assumes the values 5, 7, or 9 to accommodate varying temporal receptive fields;  $*$  symbolizes the one-dimensional convolution operator;  $b_r$  is the corresponding bias vector for that branch;  $\mu_r''$  and  $\sigma_r''$  denote, respectively, the running mean and running standard deviation maintained by the Batch Normalization (BatchNorm) layer; and  $\max(0, \cdot)$  indicates the ReLU nonlinearity, which introduces the necessary nonlinear transformations to enrich the expressive power of the learned features.

Ultimately,  $H_i$  encapsulates the multi-resolution feature set derived for the  $i$ -th charging segment. It is formed by channel-wise concatenation of the feature maps produced from  $C_i$  through the three parallel convolutional branches employing kernel sizes of  $5 \times 1$ ,  $7 \times 1$ , and  $9 \times 1$ .

Subsequently, a stack of fully connected layers arranged in an alternating Linear-ReLU-Linear-ReLU pattern is applied to  $H_i$ . This sequence performs progressive dimensionality reduction and refinement, yielding the compact, high-level deep feature representation denoted as  $F_1$ :

$$F_1 = \max(0, W'' \max(0, W' H_i + b') + b'') \quad (17)$$

Where  $W'$  and  $W''$  denote the weight matrices associated with the successive fully connected layers, while  $b'$  and  $b''$  represent the respective bias vectors for those layers.

To achieve a robust combination of time-series patterns and overall statistical summaries, the present work employs a straightforward yet effective feature fusion strategy. Specifically, the deep-learned representation  $F_1$  is merged with the normalized global statistical features from the Normalized Global Statistical Matrix (NGSM), producing the integrated, comprehensive feature vector as follows:

$$F_2 = F_1 \oplus \text{NGSM} \quad (18)$$

Where  $\oplus$  indicates the operation of feature concatenation. Consequently, the resulting integrated feature vector  $F_2$  successfully combines the time-varying dynamics captured from the sequence data with the overarching statistical properties of each charging segment. This joint representation allows the model to more holistically characterize charging behavior patterns, delivering a substantially richer and more informative input for the subsequent pattern classification step.

Following this fusion step, the enriched comprehensive feature  $F_2$  is input into a K-Means clustering algorithm to perform charging pattern categorization, thereby producing the corresponding pattern label for every charging segment:

$$\mathcal{P}_\mathcal{C} = \{\mathcal{P}_{\mathcal{C}_1}, \mathcal{P}_{\mathcal{C}_2}, \dots, \mathcal{P}_{\mathcal{C}_i}, \dots, \mathcal{P}_{\mathcal{C}_n}\} \quad (19)$$

Where  $\mathcal{P}_{\mathcal{C}_i}$  indicates the assigned charging pattern category for the  $i$ -th charging segment of Car A.

This subsection finalizes the complete pipeline—from feature engineering of the raw data through to charging pattern assignment—via four essential stages: data preprocessing (yielding NTSM and NGSM), multi-scale deep feature learning (via EV-CAE), integrated feature merging ( $F_2 = F_1 \oplus \text{NGSM}$ ), and unsupervised pattern discovery (using K-Means clustering). The combined strategy not only elevates the quality of pattern separation but also strengthens the system's robustness and versatility when confronted with intricate, heterogeneous charging datasets by effectively blending detailed temporal embeddings with broad statistical descriptors.

## Static prediction module

The static prediction module constructs a time-series forecasting model by leveraging the complete set of transformed charging segments from Car A, denoted  $\tilde{\mathcal{C}}^A = \{\mathcal{C}'_1, \mathcal{C}'_2, \dots, \mathcal{C}'_i, \dots, \mathcal{C}'_n\}$ , together with their associated pattern assignments  $\mathcal{P}\mathcal{C} = \{\mathcal{P}\mathcal{C}_1, \mathcal{P}\mathcal{C}_2, \dots, \mathcal{P}\mathcal{C}_i, \dots, \mathcal{P}\mathcal{C}_n\}$ . Here, each  $\mathcal{C}'_i$  constitutes the statistical feature sequence derived from the  $i$ -th segment, and  $\mathcal{P}\mathcal{C}_i$  specifies the pattern label allocated to that segment.

Utilizing the trained forecasting model, the module generates a static estimate for any incoming charging segment, referred to as  $\mathcal{C}^{\text{new}}$ . During operation, live measurements are gathered at regular intervals corresponding to every 2 % increase in SOC, continuing until SOC reaches 6 %. At each increment, a classification step determines the most probable pattern label; for added reliability, this label assignment may be re-evaluated up to three times. The reference charging trajectory  $\mathcal{S}$  is then obtained by computing the mean profile across all historical segments that belong to the identified pattern group, thereby producing a characteristic statistical curve representative of that behavior class. This average trajectory serves as the basis for calculating the anticipated total charging duration.

Before forecasting begins, the new segment  $\mathcal{C}^{\text{new}}$  undergoes the same statistical feature transformation process as previously described in [Eqs. 1–4](#), resulting in  $\mathcal{C}'^{\text{new}}$ . Depending on the specific use case, the core prediction architecture can be chosen flexibly from options such as LSTM, standard Recurrent Neural Network (RNN), or other time-series-appropriate designs.

In summary, the static prediction module operates as a pattern-conditioned forecasting engine, readily accommodating a variety of sequence-modeling architectures (e.g., LSTM, RNN) that can be customized to different SOC intervals, vehicle models, and operational contexts.

## Dynamic bias correction module

In the context of this study, bias correction is understood as the real-time refinement of forecast outputs by incorporating ongoing charging measurements to minimize errors introduced by external factors (such as current variations).

To counteract inaccuracies arising from such external factors (e.g., current instability), the module incorporates a collection of correction terms  $\delta = \{\delta_1, \delta_2, \dots, \delta_q, \dots\}$  that are applied dynamically to modify the baseline trajectory  $\mathcal{S}$ , thereby producing the updated, corrected trajectory  $\mathcal{S}_{\text{adjusted}}$ .

The real-time adjustment mechanism relies on the following expression for estimating the SOC value at the subsequent time step:

$$s_{q+1}^{\text{new}} = s_q^{\text{new}} + w I_q^{\text{new}} \Delta t + \delta_q \quad (20)$$

Where  $s_{q+1}^{\text{new}}$  and  $s_q^{\text{new}}$  represent the SOC values at the subsequent and present time steps, respectively, while  $I_q^{\text{new}}$  indicates the measured charging current at the current moment;  $w$  stands for the charging efficiency factor,  $\Delta t$  corresponds to the duration of the time step, and  $\delta_q$  denotes the specific bias correction term applied at iteration  $q$ .

Through repeated application of the dynamic adjustment equation outlined above, the continuously refined charging trajectory in real time can be formulated as:

$$S_{\text{adjusted}} = S + \delta \quad (21)$$

The bias correction term  $\delta$  is a pivotal element in the real-time refinement of the baseline prediction trajectory  $S$ , accounting for instantaneous charging variations.

To accommodate a range of charging conditions and varying levels of available computational resources, the present study introduces four distinct bias correction strategies, detailed as follows:

#### a. Linear Adjustment

This approach is well-suited to situations with predominantly linear deviations, applying a straightforward linear modification based on the observed discrepancies at the current step. The corresponding expression is given by:

$$\delta_q^a = k_a (I_q^{\text{new}} - \overline{I_q^{\text{new}}}) \quad (22)$$

where  $\overline{I_q^{\text{new}}}$  denotes the mean current value characteristic of the currently identified charging pattern, and  $k_a$  serves as the adjustable proportional coefficient.

This technique performs particularly well in settings with limited variability and roughly linear patterns of deviation, especially when prioritizing low computational overhead and a straightforward implementation.

#### b. Logarithmic Adjustment

This approach is especially appropriate for cases involving substantial prediction errors or pronounced deviations. It employs a logarithmic transformation to achieve smoother and more controlled corrections, with the specific expression given by the following equation:

$$\delta_q^b = \log \left( (I_q^{\text{new}} - \overline{I_q^{\text{new}}}) + k_b \right) \quad (23)$$

Where  $k_b$  acts as a tunable scaling factor that moderates the output of the logarithmic transformation, thereby controlling the degree of smoothing applied.

This technique proves especially effective for charging sequences that exhibit intermittent surges or irregular noise patterns, offering particular value in scenarios where it is critical to dampen the influence of extreme anomalies within

segments characterized by erratic or unstable current behavior.

#### c. Proportional Adjustment

This strategy performs corrections by leveraging the proportional relationship between the charging current observed at the present step and that anticipated (or measured) at the subsequent step. The mathematical expression for this correction is provided below:

$$\delta_q^c = \frac{I_{q+1}^{\text{new}} \left( I_q^{\text{new}} - \overline{I_q^{\text{new}}} \right)}{I_q^{\text{new}}} \quad (24)$$

This technique excels in environments with pronounced variability in charging dynamics, particularly in segments where the current does not remain constant and rapid shifts in current magnitude significantly influence SOC progression.

#### d. Deep Learning Adjustment

This strategy leverages a deep neural network architecture (e.g., an LSTM) to identify and model intricate nonlinear dependencies, thereby delivering highly accurate corrective adjustments. The corresponding mathematical formulation is presented below:

$$\delta_q^d = \mathcal{F} \left( t_q, I_q^{\text{new}}, s_q^{\text{new}}, o \right) \quad (25)$$

Where  $\mathcal{F}$  denotes the deep learning model's nonlinear transformation function, taking as inputs the time stamp  $t_q$ , the present current  $I_{\text{new}q}$ , the updated state-of-charge  $s_{\text{new}q}$ , together with additional feature variables  $o$ .

The proposed technique is especially appropriate for capturing intricate nonlinear charging dynamics and is best applied in high-performance contexts where modeling precision is prioritized over processing overhead.

Through the dynamic bias correction component—which merges predefined correction equations with continuously incoming measurements—this work aims to effectively compensate for forecasting inaccuracies induced by external environmental or operational factors. Simple strategies such as linear, logarithmic, and proportional scaling stand out for their minimal computational demand and very low memory footprint, making them particularly practical for real-time use cases that require fast response times under hardware constraints (limited processing capability). Conversely, whenever sufficient computing resources are available, and maximum accuracy is the primary goal, advanced neural network correction approaches (e.g., LSTM-based methods) excel at modeling complex nonlinear relationships, thereby substantially improving correction performance and producing smoother, more reliable prediction trajectories.

## Experiments

This section reports the experimental assessment of the introduced dynamic prediction scheme for electric vehicle charging time. We first describe the authentic large-scale dataset utilized throughout the research. We then examine the dynamic prediction component in detail, including an ablation analysis comparing static versus dynamic outputs with and without bias correction. The EV-CAE module's capability to identify charging behavioral patterns is confirmed through both visual inspection and numerical performance indicators. Additionally, to investigate how robustly the framework

generalizes across different neural architectures, we conduct static prediction trials using three widely used models—Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)—and compare their effectiveness. Experiments ran on a desktop workstation featuring an Intel i7–11700K processor and an NVIDIA RTX 3060 graphics card; per-vehicle training lasted between 3 and 5 minutes, and each inference step finished in no more than 2 seconds—values that support strong feasibility for online, real-time implementation.

## Dataset description and processing

The experimental results presented here rely on a collection of real operational logs from 20 electric vehicles. The data originate from the 2024 Digital Automotive Competition, hosted by the China Society of Automotive Engineers in collaboration with the National Big Data Alliance of New Energy Vehicles. An overview of the principal dataset characteristics—such as origin, temporal coverage, and file structure—appears in **Table 1**. The records encompass vital information about vehicle condition, the surrounding environment, and core charging-process metrics.

**Table 1.** Basic information about the dataset.

| Attribute              | Description                                                                                                                                                                               |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dataset name           | Electric vehicle charging and operational records                                                                                                                                         |
| Data source            | Real-world operational and charging logs collected from 20 electric vehicles during 2023 and 2024                                                                                         |
| Data scale             | Second-by-second timestamped measurements available for each vehicle across both 2023 and 2024                                                                                            |
| Sample format          | Tabular structure where every row corresponds to one timestamped observation, encompassing vehicle state variables, surrounding environmental conditions, and charging process parameters |
| Recording range        | From April 2023 through April 2024                                                                                                                                                        |
| Data storage method    | Yearly separation: data for each vehicle is organized into distinct files—one dedicated to 2023 and another to 2024                                                                       |
| Total number of fields | 35 attributes in total                                                                                                                                                                    |

Charging and driving activities of the 20 EVs were logged continuously from 2023 through 2024. Records for each vehicle are organized into separate yearly files; for example, 30 days of uninterrupted measurements for one vehicle occupy a single file. Sampling occurs every second and includes parameters related to driving state, charging state, ambient conditions, and battery properties. Because raw acquisition occasionally introduced gaps, repeated entries, or outlier values, dedicated cleaning and refinement procedures were essential to reach acceptable quality levels. The unprocessed collection comprises both driving episodes and charging events across all 20 vehicles. Creation of a clean, study-appropriate subset involved three successive operations: chronological splitting of the timeline, removal or correction of problematic records, and final subdivision of the cleaned material.

Once cleaned, the 2023 portion showed an average of 78.55 charging events per vehicle and an average accumulated charging time of 215.45 hours per vehicle. In 2024, the mean count of charging events rose to 84.25 per vehicle, while the mean total charging time fell to 200.45 hours. Per-vehicle statistics on charging-event counts and durations are summarized in **Table 2**. In all subsequent tests, the full 2023 records serve for pattern discovery and model fitting, while the 2024 records are held out strictly for evaluation. Maintaining complete separation between the two calendar years guarantees unbiased testing and provides a realistic measure of the model’s ability to generalize across time.

**Table 2.** Overview of vehicle charging and driving segments after data processing.

| Vehicle ID | Charging sessions (2023 / 2024) | Total charging duration (hours) (2023 / 2024) | Driving sessions (2023 / 2024) | Total driving duration (hours) (2023 / 2024) |
|------------|---------------------------------|-----------------------------------------------|--------------------------------|----------------------------------------------|
| 01         | 92 / 99                         | 464 / 446                                     | 1230 / 451                     | 248 / 310                                    |
| 02         | 162 / 184                       | 500 / 463                                     | 875 / 1330                     | 961 / 1007                                   |
| 03         | 66 / 49                         | 284 / 65                                      | 463 / 462                      | 372 / 426                                    |
| 04         | 41 / 45                         | 34 / 50                                       | 462 / 1188                     | 462 / 886                                    |
| 05         | 84 / 52                         | 319 / 227                                     | 315 / 171                      | 195 / 104                                    |
| 06         | 21 / 60                         | 199 / 352                                     | 339 / 539                      | 213 / 340                                    |
| 07         | 50 / 39                         | 159 / 110                                     | 284 / 308                      | 155 / 155                                    |
| 08         | 25 / 8                          | 63 / 22                                       | 209 / 183                      | 132 / 108                                    |
| 09         | 96 / 138                        | 91 / 168                                      | 908 / 1258                     | 691 / 764                                    |
| 10         | 126 / 58                        | 343 / 195                                     | 337 / 399                      | 626 / 583                                    |
| 11         | 36 / 46                         | 29 / 42                                       | 241 / 291                      | 175 / 186                                    |
| 12         | 88 / 116                        | 147 / 135                                     | 544 / 501                      | 407 / 296                                    |
| 13         | 49 / 114                        | 124 / 260                                     | 592 / 354                      | 341 / 241                                    |
| 14         | 129 / 131                       | 427 / 363                                     | 462 / 272                      | 364 / 281                                    |
| 15         | 1 / 104                         | 1 / 240                                       | 2318 / 502                     | 271 / 316                                    |
| 16         | 84 / 53                         | 179 / 93                                      | 247 / 222                      | 583 / 461                                    |
| 17         | 53 / 110                        | 93 / 181                                      | 395 / 423                      | 262 / 306                                    |
| 18         | 113 / 109                       | 389 / 320                                     | 380 / 259                      | 636 / 248                                    |
| 19         | 232 / 135                       | 438 / 239                                     | 636 / 348                      | 711 / 233                                    |
| 20         | 23 / 35                         | 26 / 38                                       | 227 / 194                      | 115 / 108                                    |

### Experimental validation of dynamic charging duration prediction

Here, we demonstrate the practical value of the suggested dynamic charging-time forecasting framework. A side-by-side comparison of purely static estimates with dynamically adjusted ones clearly shows that continuous incorporation of fresh measurements yields markedly superior accuracy.

Consider Car A as a typical instance (experiments treat each of the 20 vehicles independently). The task is to anticipate the remaining charging time for every charging event that took place during 2024, using only the complete 2023 charging history of that vehicle (current profiles, starting SOC values, temperature readings, etc.) plus a very short prefix—e.g., the opening 5 minutes—of the target 2024 charging session.

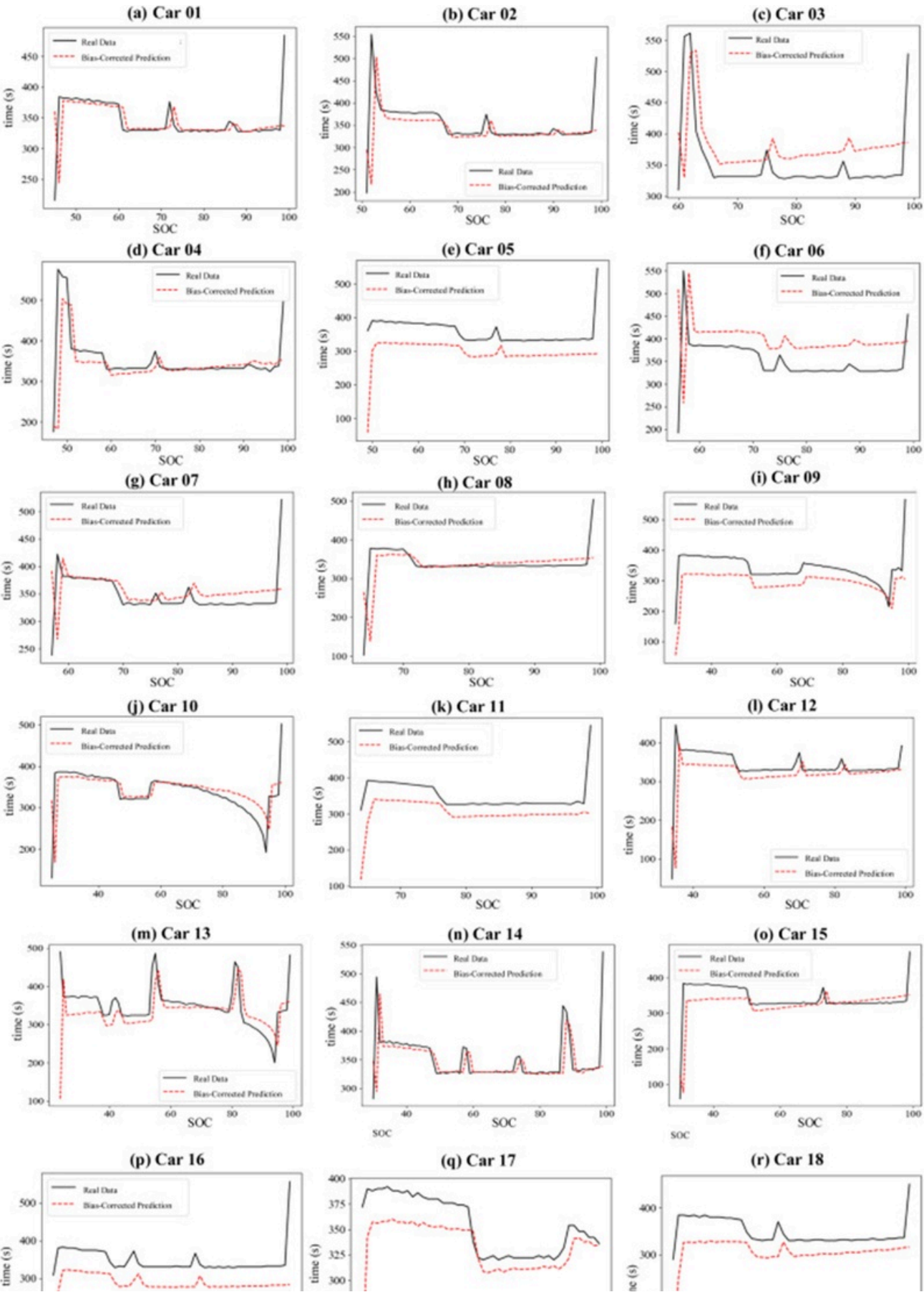
The overall prediction pipeline follows a data-driven sequence consisting of four key phases: data preprocessing, pattern-based clustering of charging behavior, initial static estimation, and ongoing dynamic refinement. The process starts by feeding Car A's 2023 charging records into the preprocessing block illustrated in [Figure 1](#). The cleaned segments are then passed to the pattern classification block, which groups them into meaningful charging archetypes.

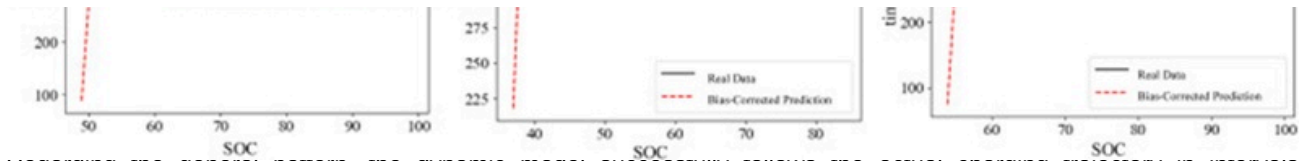
These labeled 2023 segments form the training corpus. For prediction, only the first 5 minutes of a given 2024 charging event are supplied to the static prediction stage, which—guided by the learned pattern labels—delivers an initial full-session duration estimate that holds independently of the observed starting SOC.

The static estimates subsequently enter the dynamic adjustment stage. There, live incoming data from the same 2024 charging event are used to apply progressive corrections, enabling the system to revise its forecast at each future time step and ultimately deliver highly refined dynamic predictions.

### Comparison of dynamic prediction curves with actual outcomes

**Figure 3** illustrates the dynamic prediction trajectories (red dashed lines) for 18 vehicles alongside the observed charging processes (black solid lines), with linear bias correction. The x-axis ranges from 0 % to 100 % in 1 % increments, and the y-axis shows the time (in seconds) required to advance from the current SOC to the next 1 % increment. Visual inspection reveals how the dynamic forecasts progressively refine themselves to align more closely with the true charging behavior over time.





Regarding the general pattern, the dynamic model successfully follows the actual charging trajectory in intervals characterized by gradual SOC progression. In these regions, the predicted and observed curves exhibit strong agreement, reflecting excellent predictive precision. Nevertheless, noticeable discrepancies emerge during phases of rapid SOC variation (for instance, in subplots **Figure 3(m)** for Car 13 and **Figure 3(p)** for Car 16). Such deviations likely stem from sparse historical charging records, resulting in incomplete coverage during charging pattern classification. Specifically, Car 13 recorded only 49 charging events in 2023, and Car 16 recorded only 35—both substantially lower than the yearly average of 78.55 sessions. These cases demonstrate that the proposed bias-correction strategy maintains robust adaptive correction even with limited training data.

Furthermore, model fit was quantified using the coefficient of determination ( $R^2$ ), which ranges from 0 to 1; values closer to 1 indicate a closer correspondence between predicted and actual charging profiles. Across all 2024 charging segments for the 20 vehicles, the dynamic charging duration prediction framework was applied with each bias correction strategy, producing corresponding dynamic forecasts. The evaluation yielded  $R^2$  values of 0.95 for linear bias correction, 0.948 for logarithmic bias correction, 0.952 for proportional bias correction, and 0.96 for the deep learning-based bias correction approach.

All four correction techniques delivered  $R^2$  scores exceeding 0.94, confirming consistently high performance. The logarithmic method showed the lowest fit ( $R^2 = 0.948$ ), whereas the deep learning method delivered the highest ( $R^2 = 0.96$ ). These outcomes reinforce the overall strength of the proposed dynamic prediction framework. They also indicate that, although deep learning corrections excel in data-rich settings, simpler techniques such as linear or proportional adjustments remain highly reliable while requiring far less computational effort.

### Comparison of static versus dynamic prediction outcomes

To confirm the value added by the dynamic bias correction component within the overall prediction framework, this study compares the initial charging curves generated by the static prediction module with the refined curves resulting from dynamic adjustments. **Figure 4** highlights the behavior of one representative vehicle across various prediction phases. **Figure 4(a)** depicts the real-time evolution of charging current throughout an entire session. At the same time, **Figure 4(b)** compares the actual charging profile against both the uncorrected static forecast and the bias-corrected dynamic forecast.

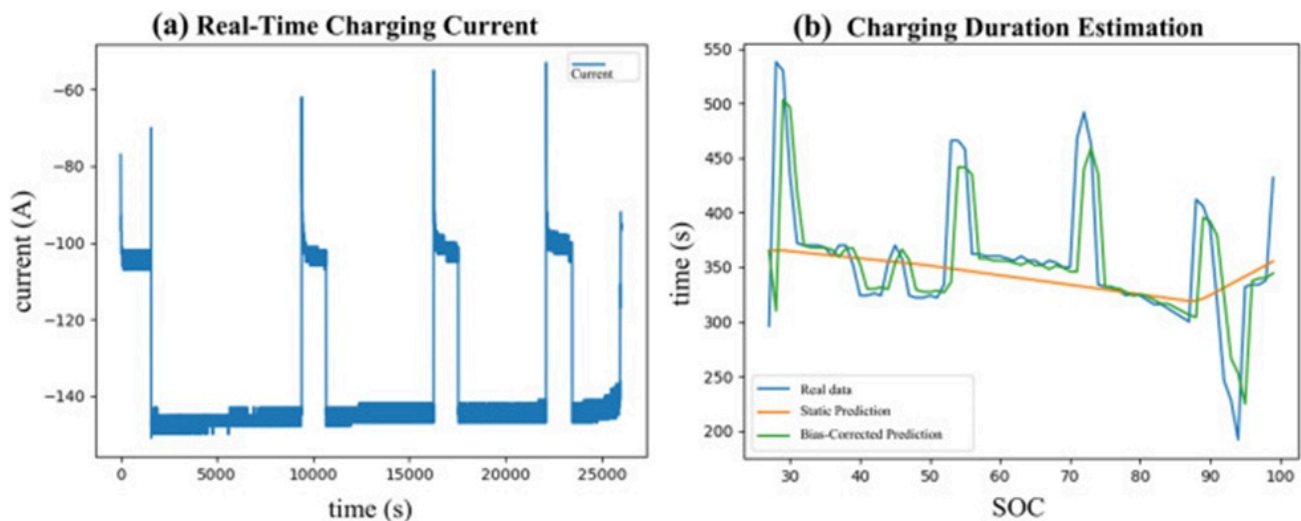


**Figure 4.** Charging duration forecasting for an individual vehicle. (a) Time series of real-time charging current over a complete session (x-axis: time (s); y-axis: current (A)). (b) Overlay of actual measurements (blue), static prediction output (orange), and dynamic bias-corrected prediction (green) as functions of SOC (%). The application of dynamic correction visibly enhances conformity with the observed charging data.

From **Figure 4(b)**, it is clear that the uncorrected predicted trajectory (orange line) generally captures the overall shape of the true charging curve (blue line). However, substantial mismatches occur in certain phases, especially during rapid SOC transitions, when the model fails to respond promptly to behavioral shifts. Once bias correction is incorporated, the system performs incremental refinements at every time step by leveraging live charging measurements, resulting in the corrected forecast (green line) adhering much more closely to the observed data and markedly reducing discrepancies in volatile segments. When viewed alongside the real-time current profile in **Figure 4(a)**, this highlights the particular strength of the bias adjustment approach during episodes of pronounced current swings, enabling rapid detection and accommodation of evolving charging modes. Although detailed per-SOC-interval error quantification was not conducted here, insights from **Figure 3** and **Figure 4** indicate that the static prediction stage delivers strong alignment in zones of gradual SOC progression. In contrast, the dynamic correction stage excels in zones with abrupt SOC shifts (e.g., 70 %–90 %). Collectively, these observations underscore the framework’s robust flexibility and consistency under a wide range of SOC dynamics.

Additionally, this work employs mean squared error (MSE) and  $R^2$  as metrics to assess the outputs of the static prediction module. MSE quantifies the average squared deviation and serves as an indicator of prediction precision; its possible values span  $[0, \infty]$ , with values approaching 0 denoting superior accuracy in estimated charging times. For every charging event across the 20 vehicles in 2024, MSE was computed by contrasting the static forecasts with the corresponding actual durations, then averaging the results across all vehicles and sessions to gauge global predictive error. The  $R^2$  computation follows the same procedure. It is important to emphasize that dynamic prediction outcomes incorporate ongoing real-time observations from the target segments themselves, rendering direct MSE comparisons for those results inappropriate.

Experimental findings reveal that static prediction yields an MSE of 0.452 and an  $R^2$  of 0.948. These figures demonstrate that, even absent dynamic bias correction, the static module attains considerable precision in forecasting charging durations and provides a reliable representation of the charging trajectory. This outcome further substantiates the efficacy of the charging pattern classification stage in separating distinct behavioral categories, thereby providing a reliable basis for subsequent forecasting steps. To quantify the incremental benefit of each bias-correction variant, an ablation study was conducted. As detailed in **Table 3**, all four correction approaches outperform the static baseline in  $R^2$ , confirming the value of dynamic refinement. The proportional and deep learning-based methods deliver the strongest accuracy gains, with the deep learning variant reaching the peak overall score ( $R^2 = 0.960$ ). Such results suggest that lightweight methods, such as linear correction, deliver efficient improvements at minimal cost. In contrast, deep learning correction proves optimal in scenarios with ample data volume and processing capacity, where top-tier precision is prioritized.

**Table 3.** Ablation comparison of bias correction strategies.

| Method                   | Bias correction applied | $R^2$ score |
|--------------------------|-------------------------|-------------|
| Static prediction only   | ✗                       | 0.948       |
| Linear correction        | ✓                       | 0.95        |
| Logarithmic correction   | ✓                       | 0.948       |
| Proportional correction  | ✓                       | 0.952       |
| Deep learning correction | ✓                       | 0.96        |

Relative to static forecasts, dynamic versions consistently outperform. Within the dynamic bias correction component, all four adjustment techniques yield  $R^2$  scores higher than those of the static module. This superiority arises because dynamic prediction revises its estimate iteratively at each timestep using fresh incoming data, thereby achieving tighter synchronization with ground-truth measurements. Consequently, the inclusion of bias adjustment mechanisms confers a clear  $R^2$  advantage to dynamic predictions over static ones, reinforcing the mechanism's role in elevating tracking fidelity and curve-matching quality throughout the charging sequence.

Compared with approaches that rely on charging-pile-side measurements [3], the dynamic prediction framework introduced here draws exclusively from on-vehicle BMS signals. This choice facilitates more granular monitoring of internal battery states and nuanced charging variations. The vehicle-centric data strategy broadens practical utility, especially in edge computing setups or areas with limited infrastructure, where detailed station-level or grid-side records may be inaccessible.

Regarding generalization capability, the method was rigorously tested on 20 diverse real-world electric vehicles exhibiting heterogeneous charging profiles. By contrast, prior studies such as [9] relied on static models in controlled settings and with restricted sample sizes (e.g., 4 vehicles), limiting their versatility. Moreover, unlike purely static machine learning techniques exemplified by [22] that offer no real-time adaptation, the present framework embeds a versatile multi-strategy bias correction layer (encompassing linear, logarithmic, proportional, and deep learning-based options) capable of continuously updating forecasts timestep by timestep. This architecture markedly improves concordance between predicted and observed charging paths, particularly amid noisy or highly variable conditions.

Although sophisticated deep learning architectures, including graph-based models [3], have shown encouraging results for station- or region-scale forecasting, they generally require substantial computational resources and vast spatiotemporal datasets. In opposition, the current framework preserves modest resource requirements and supports genuine real-time vehicle-level execution without compromising forecast quality (attaining  $R^2$  values as high as 0.96).

Overall, the presented framework effectively harmonizes predictive precision, deployment practicality, and temporal responsiveness. It delivers a flexible, implementable approach to EV charging duration estimation suitable for varied real-world settings. That said, its resilience in highly irregular or outlier-heavy circumstances is recognized as a constraint. Prospective research could incorporate anomaly-detection routines and adaptive data-refinement steps to enhance robustness in challenging edge cases.

## Determination of cluster count and assessment within K-Means

The present subsection examines one of the pivotal decisions in charging behavior grouping: identifying and validating the most suitable number of clusters for K-Means. An appropriate choice of the number of clusters is vital, since it lays a solid, consistent foundation for accurate pattern recognition, which, in turn, supports effective feature learning and reliable downstream forecasting.

This investigation selects the optimal cluster number for each vehicle by using two established performance indicators: Sum of Squared Errors (SSE) and Silhouette Coefficient. These metrics together guarantee trustworthy pattern partitioning outcomes. SSE measures how tightly the members of a cluster are packed; progressively smaller SSE values indicate greater homogeneity within clusters. The Silhouette Coefficient jointly assesses internal cohesion and external separation, with scores nearing 1 reflecting well-defined, distinct groups. Combining evidence from both measures enables a reasoned, data-driven selection of the optimal cluster cardinality.

To provide more direct insight into how clustering behaves, t-SNE is used to compress the high-dimensional feature space into 2D, enabling visual inspection of cluster arrangements for different  $k$  values (Figures 5 and 6). The resulting plots and metrics jointly confirm the soundness of the adopted clustering workflow. Figure 5 traces SSE evolution across a range of cluster counts for various vehicles, and Figure 6 tracks corresponding Silhouette Coefficient trends. Both

figures clearly show that applying EV-CAE feature transformation beforehand leads to markedly tighter clusters and better separation between them. **Table 4** summarizes the final optimal cluster numbers per vehicle obtained using the combined EV-CAE + K-Means procedure. These per-vehicle decisions directly inform the pattern-based charging time predictions and analyses that follow.

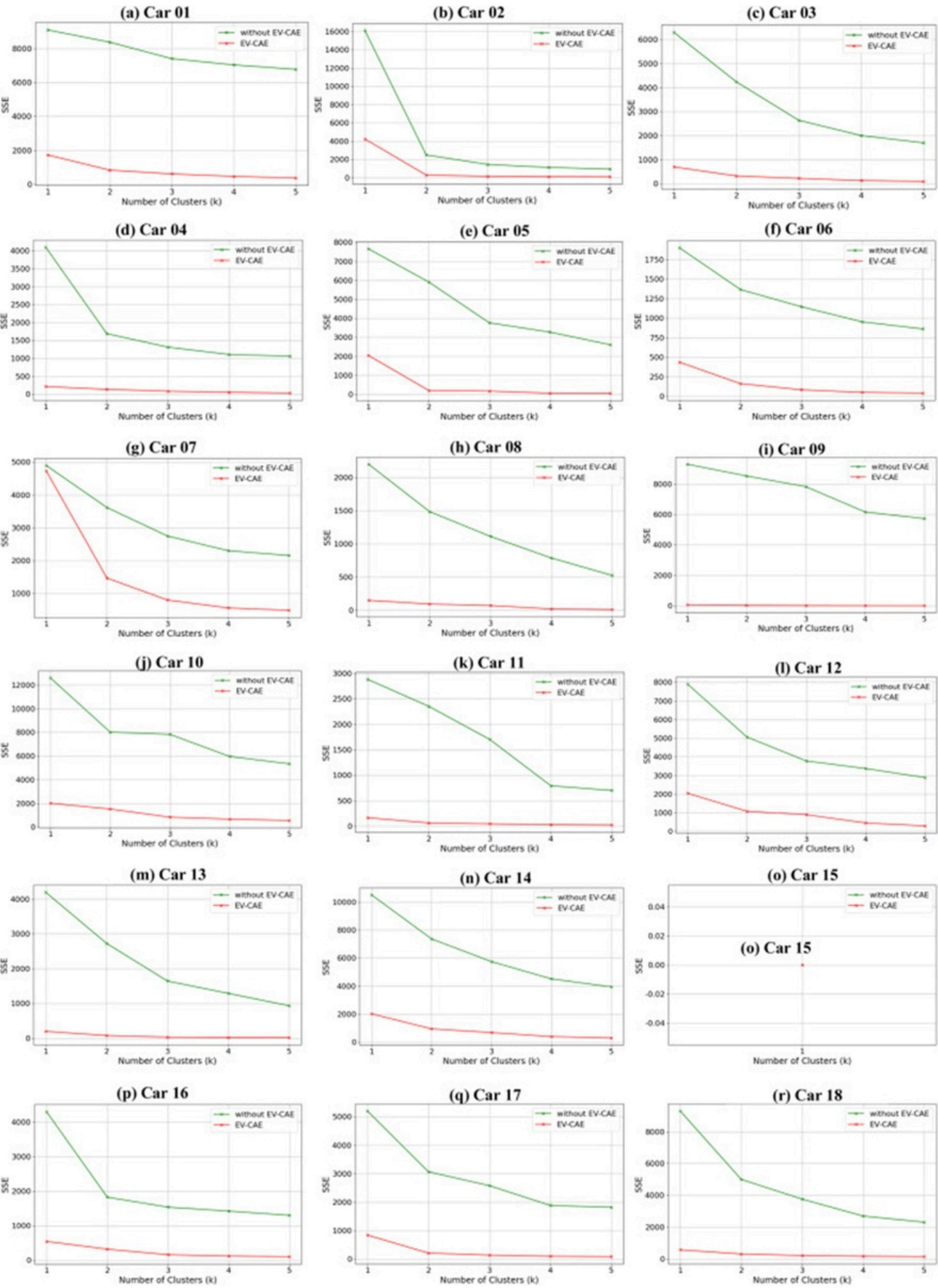


**Figure 5.** Comparison of clustering performance for 18 electric vehicles. Subplots (a)–(r) plot the sum of squared errors (SSE) values with EV-CAE preprocessing (red line) and without it (green line) as a function of cluster count. The x-axis is the number of clusters  $k$ ; the y-axis is SSE. Smaller SSE values correspond to more compact clusters.

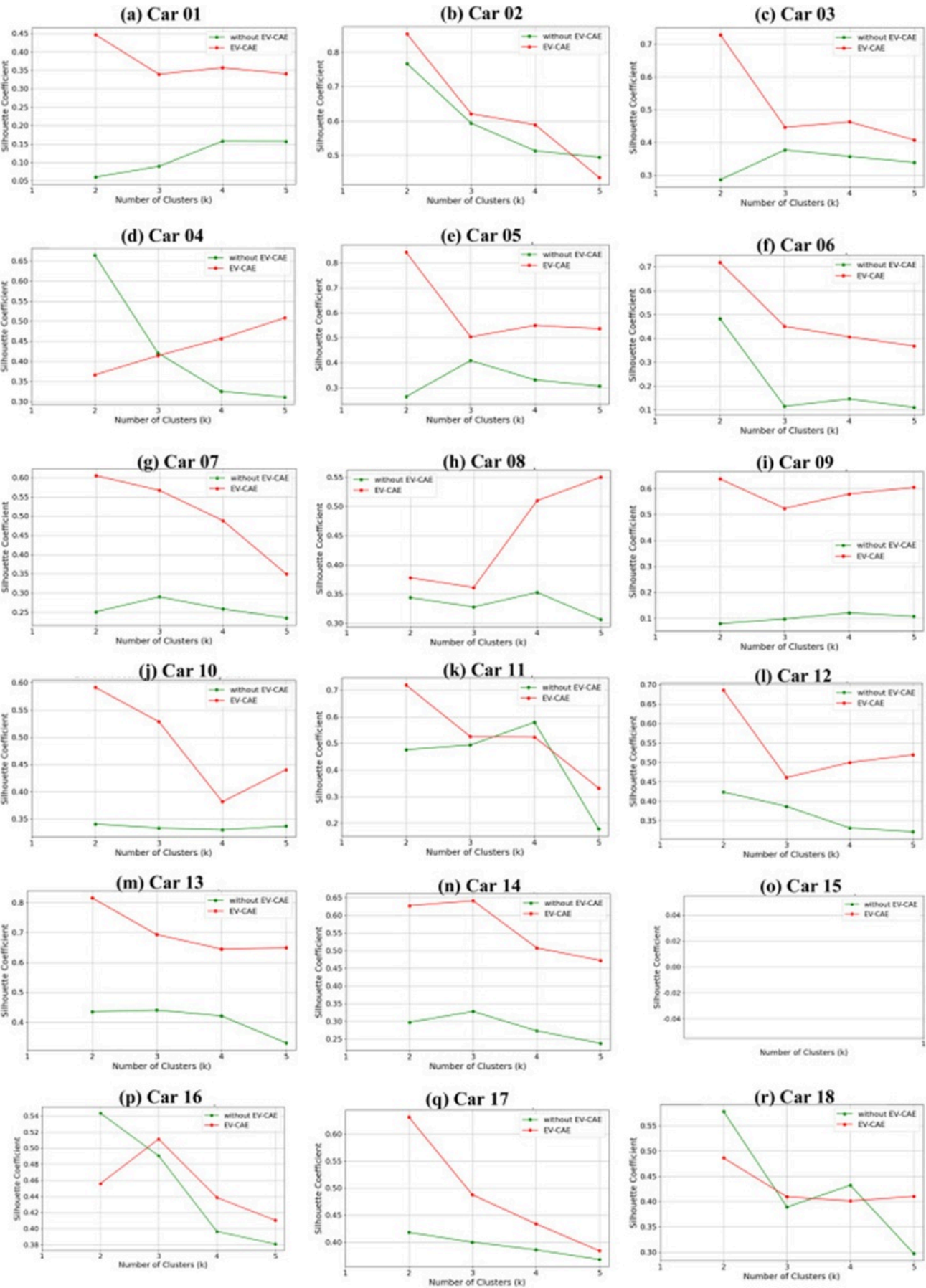


**Figure 6.** Assessment of clustering quality for 18 electric vehicles based on the silhouette coefficient. Subplots (a)–(r) display silhouette scores with EV-CAE feature extraction (red) and without (green) across varying cluster counts. The x-axis indicates k (number of clusters), while the y-axis shows the silhouette coefficient. Higher scores denote clearer cluster boundaries.

**Table 4.** Optimal number of clusters for each vehicle using the EV-CAE and K-Means method.

| Car | Cluster | Car | Cluster | Car | Cluster | Car | Cluster |
|-----|---------|-----|---------|-----|---------|-----|---------|
| 1   | 2       | 6   | 2       | 11  | 2       | 16  | 3       |
| 2   | 2       | 7   | 2       | 12  | 2       | 17  | 2       |
| 3   | 2       | 8   | 4       | 13  | 2       | 18  | 2       |
| 4   | 2       | 9   | 2       | 14  | 2       | 19  | 2       |
| 5   | 2       | 10  | 2       | 15  | 1       | 20  | 3       |

Confirmation of EV-CAE effectiveness for feature representation and forecasting

This segment supplies further evidence supporting the role of EV-CAE as a powerful feature extractor and contributor to prediction quality. Acting as an integral part of the dynamic forecasting pipeline, EV-CAE simultaneously refines the precision of pattern categorization and furnishes richer, more informative representations that strengthen both the static baseline and the subsequent dynamic refinement stages. Visual demonstrations via t-SNE, together with quantitative benchmarks using an LSTM forecaster, show that the inclusion of EV-CAE substantially improves the quality of both grouping and time-to-charge estimation.

Visual inspection of results

Visualization serves here as a primary means of verifying EV-CAE's added value. The pattern classification block ultimately assigns each charging session to a given vehicle. To clearly demonstrate the enhancement provided by EV-CAE during representation learning, t-SNE is used to reduce feature dimensionality to 2D, enabling direct visual comparison between two setups: standard K-Means applied to the original raw features versus K-Means applied after the EV-CAE transformation.

As shown in **Figure 7(a)**, the cluster layout obtained by directly applying K-Means clustering to unprocessed charging records (without EV-CAE involvement) exhibits considerable spread and heavy overlap across the supposed categories. By comparison, **Figure 7(b)** shows the outcome when EV-CAE extracts features before clustering; the points form much more distinct groups with far less inter-category mixing. The side-by-side views make it apparent that EV-CAE generates features possessing greater discriminative power, which translates into sharper, more stable, and ultimately more accurate assignment of charging patterns.

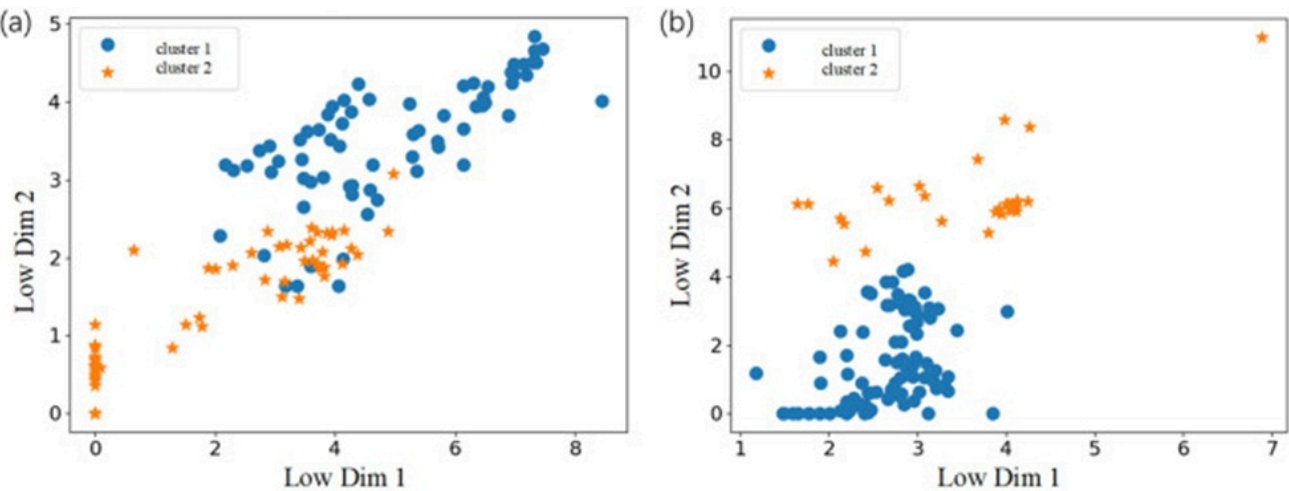


**Figure 7.** Two-dimensional t-SNE projections of clustering outcomes for Car 10 obtained via K-Means: (a) Result when clustering is performed on the original raw features (no EV-CAE), (b) The result after feature extraction through the EV-CAE module precedes K-Means. In both panels, axes represent the abstract low-dimensional coordinates produced by t-SNE (“Low Dim 1” and “Low Dim 2”).

Quantitative evaluation

To more rigorously measure the contribution of EV-CAE to feature quality, this work applies the EV-CAE combined with the K-Means procedure to assign charging pattern labels separately for the 2023 and 2024 datasets, ensuring complete independence between the two years and preventing information leakage. In the setup, the 2023 charging segments, together with their derived pattern labels, serve as the training corpus for an LSTM classifier. The 2024 segments serve as the unseen test set for evaluating the trained model’s classification performance. Classification performance is then quantified by comparing the LSTM’s predicted pattern labels against the ground-truth labels (obtained from the same EV-CAE+K-Means process on 2024 data) for every vehicle.

To demonstrate EV-CAE’s superiority more clearly, its results are benchmarked against a straightforward, manual, rule-based segmentation approach. In the manual method, charging patterns are divided into four groups based on average current levels: Pattern 1 for currents  $\leq 200$  A, Pattern 2 for 200–600 A, Pattern 3 for 600–1000 A, and Pattern 4 for  $> 1000$  A. The complete set of manually assigned pattern labels across all vehicles appears in [Table 5](#).

**Table 5.** Charging pattern classification based on manual segmentation rules.

| Car number | Cluster | Car number | Cluster | Car number | Cluster | Car number | Cluster |
|------------|---------|------------|---------|------------|---------|------------|---------|
| 1          | 1       | 6          | 1       | 11         | 4       | 16         | 3       |
| 2          | 4       | 7          | 3       | 12         | 3       | 17         | 3       |
| 3          | 3       | 8          | 2       | 13         | 2       | 18         | 3       |
| 4          | 4       | 9          | 3       | 14         | 1       | 19         | 4       |
| 5          | 3       | 10         | 3       | 15         | 1       | 20         | 3       |

**Table 6** summarizes the per-vehicle classification accuracies achieved by the LSTM model when trained and tested under the two competing approaches: EV-CAE+K-Means versus manual segmentation. Entries marked with “/” indicate vehicles that exhibit only a single charging pattern under the respective method, rendering accuracy computation undefined. The figures show that the EV-CAE+K-Means pipeline markedly improves pattern recognition performance,

yielding an average accuracy of 89.33% with the LSTM classifier. In particular, EV-CAE proves highly effective at handling intricate and variable charging behaviors, displaying excellent generalization and robustness across diverse scenarios.

**Table 6.** LSTM prediction accuracy for EV-CAE+K-Means and manual segmentation.

| Car number | EV-CAE + K-Means | Manual segmentation | Car number | EV-CAE + K-Means | Manual segmentation |
|------------|------------------|---------------------|------------|------------------|---------------------|
| 1          | 0.8687           | /                   | 11         | 0.8537           | 0.6829              |
| 2          | 0.9945           | 0.8729              | 12         | 0.9266           | 0.789               |
| 3          | 0.9556           | 0.7333              | 13         | 0.9099           | 0.933               |
| 4          | 0.9268           | 0.6585              | 14         | 0.7568           | /                   |
| 5          | 0.9583           | 0.9167              | 15         | /                | /                   |
| 6          | 0.8846           | /                   | 16         | 0.9324           | 0.8784              |
| 7          | 0.8378           | 0.973               | 17         | 0.8095           | 0.7038              |
| 8          | 0.6667           | 0.83/1              | 18         | 0.999            | 0.7848              |
| 9          | 0.9697           | 0.2879              | 19         | 0.938            | 0.8272              |
| 10         | 0.9787           | 0.5745              | 20         | 0.8065           | 0.5161              |

By contrast, the manual segmentation baseline yields an average LSTM accuracy of only 69.78%, with the poorest individual-vehicle performance as low as 28.79%—substantially below the EV-CAE+K-Means figures. This gap highlights the limitations of fixed-threshold rules when faced with complex, irregular, or outlier charging profiles. Across the board, the EV-CAE+K-Means combination delivers consistently higher accuracy across different vehicles, confirming its robust adaptability to a broad spectrum of real-world charging dynamics.

Validation of the generalization capability of the static prediction module

The current section evaluates how well the static prediction module accommodates different neural architectures by benchmarking three representative models: Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM). To guarantee equitable and reproducible comparisons, identical experimental conditions are enforced for all architectures: (1) an identical input feature set comprising SOC, voltage, current, and the charging pattern label produced by EV-CAE; (2) uniform training hyperparameters, including the Adam optimizer, learning rate of 0.001, batch size of 64, and 100 training epochs. These controls isolate performance variations to differences in model architecture rather than training protocol discrepancies.

The static prediction module is intentionally designed to be compatible with a variety of sequential or time-aware models, reflecting its inherent flexibility. To substantiate this property, the present experiments evaluate the forecasting quality of MLPs, RNNs, and LSTM using the unified 20-vehicle dataset. Model effectiveness is quantified via two standard regression metrics:  $R^2$  and MSE.

**Figure 8** illustrates the comparative results of the three models within the static prediction framework. Among them, RNN records the highest mean  $R^2$  of 95.5%, outperforming MLP (the lowest performer) by roughly 0.74%. In terms of MSE, LSTM achieves the smallest error, while RNN exhibits the largest, with a relative gap of 5.69%. These patterns are consistent with the intrinsic strengths and limitations of each architecture: MLP’s feedforward design excels at nonlinear mappings but disregards temporal structure; RNN effectively captures sequential dependencies yet remains vulnerable to

vanishing gradients; LSTM, thanks to its gating mechanisms, handles extended temporal dependencies more reliably, which accounts for its edge in minimizing squared error.

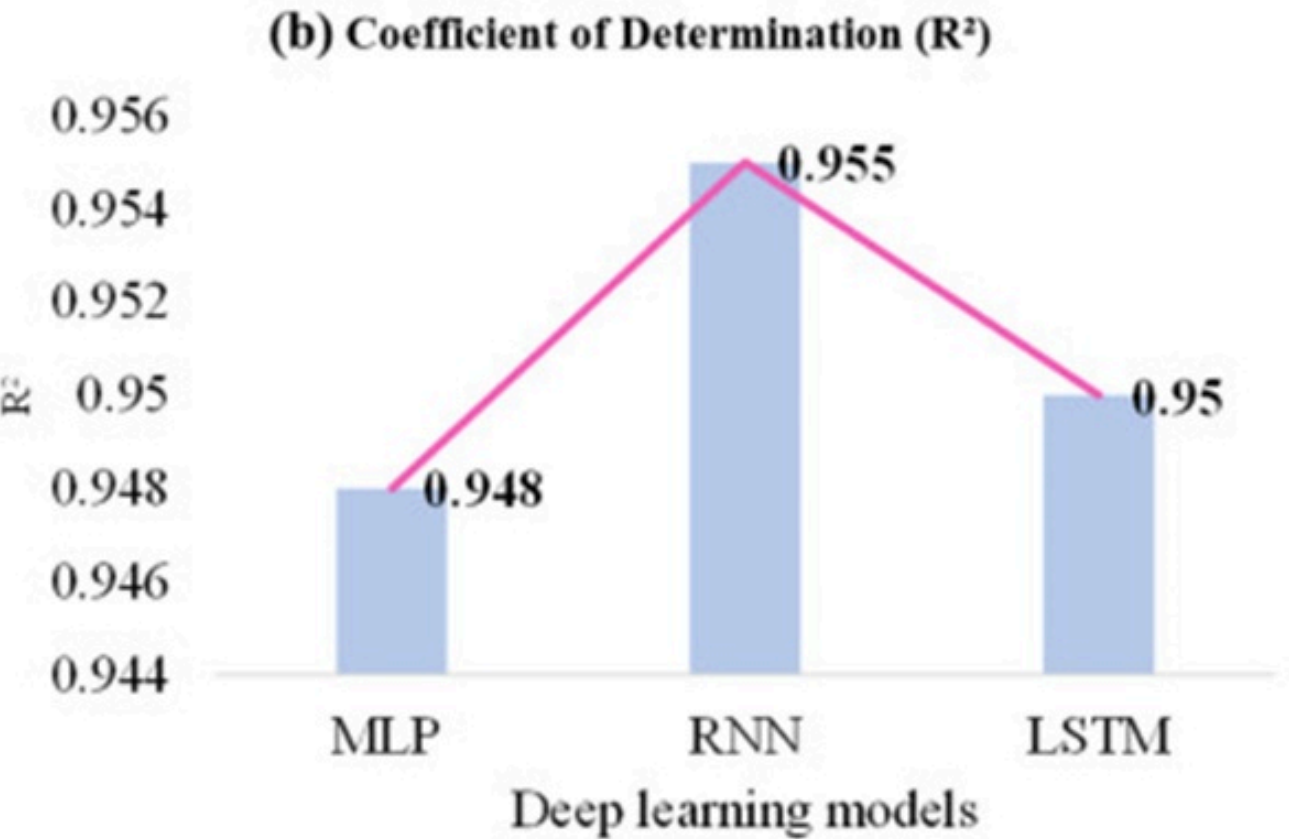
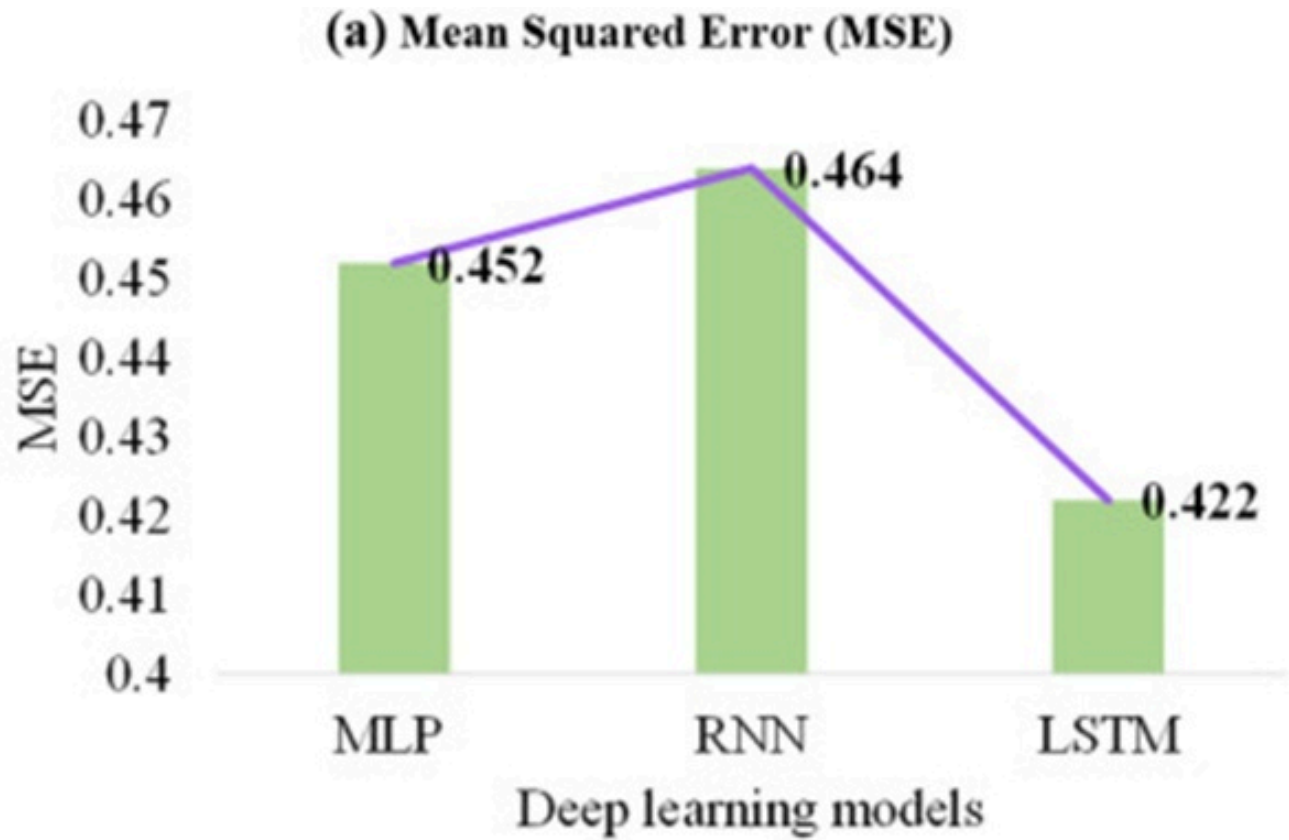


**Figure 8.** Performance comparison of three deep learning models in the static prediction module: (a) Mean Squared Error (MSE); (b) Coefficient of Determination ( $R^2$ ). The x-axis denotes the model type (MLP, RNN, LSTM), and the y-axis shows the corresponding evaluation metric.

Taken together, the experiments reveal that MLP, RNN, and LSTM all deliver strong predictive accuracy when embedded in the static prediction module. The modest spread in  $R^2$  and MSE values across the three architectures further confirms the module's high degree of model-agnostic adaptability, ensuring dependable, high-quality charging duration estimates regardless of the chosen deep learning backbone.

## Conclusion

This research introduces a comprehensive, integrated framework designed for predicting electric vehicle (EV) charging durations. It combines unsupervised identification of charging patterns, static baseline forecasting, and a suite of dynamic bias correction techniques. The approach successfully addresses major obstacles, including the wide variability in charging behavior, the demand for real-time vehicle-side responsiveness, and constraints on data availability.

The framework operates in successive stages. Initially, an improved convolutional autoencoder (EV-CAE) captures multi-resolution features from charging records, followed by K-Means clustering that automatically partitions charging behaviors into distinct groups without relying on manual labels or fixed thresholds, thereby exhibiting robust generalization. In the subsequent stage, a static prediction component generates an initial estimate of charging time using inputs such as SOC, voltage, current, and the derived pattern labels. To achieve greater sensitivity to evolving conditions and higher precision, a dynamic bias correction layer is incorporated. This layer applies one of four correction strategies—linear, logarithmic, proportional, or deep learning-based—to iteratively refine the forecast at each time step using incoming real-time charging measurements.

Validation experiments conducted on operational data from 20 actual electric vehicles confirm that the pattern recognition stage substantially boosts accuracy during static prediction. In the dynamic phase, activating bias correction increases the  $R^2$  metric from 0.948 to 0.960, markedly improving temporal synchronization and the framework's ability to track irregular variations. Moreover, the model exhibits consistent reliability even in low-temperature environments (with an average ambient temperature of 5 °C), underscoring its resilience to environmental factors and its readiness for practical deployment.

While the method performs robustly under standard operating scenarios, future efforts should aim to broaden the dataset by including vehicles from additional geographic regions, climatic zones, and model types. Furthermore, linking charging duration forecasts with energy management and scheduling algorithms could enhance the framework's utility in smart transportation ecosystems, supporting both efficient charging coordination and active participation in electricity markets.

### (1) Segmentation Based on the Timeline

In the original dataset, yearly records for each vehicle are stored in a single continuous file. Nevertheless, temporal interruptions frequently occur—for example, sampling may shift from 10-second or 2-second intervals to a gap of 8 hours. These discontinuities can compromise the validity of later analytical steps and reduce the reliability.

To overcome this limitation, the data are divided into temporally continuous blocks. Each block corresponds to an uninterrupted operational period encompassing both driving and charging activities. In the present work, segmentation uses a maximum allowable gap of 60 minutes between successive records, thereby preserving chronological coherence within every resulting segment.

### (2) Data Cleaning

Data refinement primarily targets four categories of defects: duplicated entries, absent values, timestamp disorders (such as reversed chronological order), and anomalous measurements. An overview of the cleaning operations and remedies is provided in **Table 7**. Conventional techniques adequately resolve duplicates, missing entries, and ordering errors, whereas outlier management demands more sophisticated handling.

Additional specifics on handling missing data and outliers follow. Missing entries predominantly affect fields such as current, voltage, and SOC, yet they constitute less than 1% of the observations within any given charging segment (sampled at 2-second intervals). Due to their sparse, seemingly random occurrence, these gaps are discarded to prevent the introduction of artificial variability into high-frequency time series. Outlier identification and mitigation rely on three complementary rules:

- a. Charging status anomalies: Isolated erroneous points are attenuated via modal smoothing applied over a sliding temporal window.
- b. Physically impossible readings: Measurements that breach fundamental physical limits (for instance, SOC exceeding 100 %) are adjusted through interpolation from neighboring valid points.
- c. Sudden statistical discontinuities: Abrupt shifts are corrected by enforcing local statistical coherence.

This multi-criteria filtering safeguards data quality and minimizes the adverse impact of irregularities on subsequent modeling procedures.

**Table 7.** Common data quality issues and corresponding remedies.

| Type             | Action                                                                             |
|------------------|------------------------------------------------------------------------------------|
| Duplicate Data   | Remove duplicates.                                                                 |
| Missing Data     | Directly delete rows containing missing values.                                    |
| Data Sorting     | Sort the data in chronological order.                                              |
| Outlier Handling | Apply rule-based filtering: smoothing, boundary check, and statistical adjustment. |

(3) Data Partitioning

The cleaned dataset is subsequently subdivided to isolate charging episodes from driving episodes. The analyses and models developed in this study concentrate exclusively on charging-related data for experimental purposes.

(1) Charging Data Extraction Criteria:

- a. Charging status indicator: The flag must denote “charging in progress.”
- b. Current value: Current readings are negative, signifying energy inflow from an external charger.
- c. Duration: Each extracted charging segment must span at least 30 minutes.

(2) Driving Data Extraction Criteria:

- a. Vehicle status: The status flag indicates “active.”

- b. Speed: Vehicle speed is  $\geq 0$ .
- c. Zero-speed duration: Periods of zero speed must not exceed 10 minutes.

These explicit rules enable reliable differentiation between charging and driving phases, producing a coherent, focused dataset well-suited for investigating charging dynamics.

## Abbreviations

BMS: Battery Management System

CC: Constant Current

CNN: Convolutional Neural Networks

CV: Constant Voltage

EV: Electric Vehicle

EV-CAE: Electric Vehicle Charging Convolutional Autoencoder

IEA: International Energy Agency

LSTM: Long Short-term Memory

MSE: Mean Squared Error

MLP: Multilayer Perceptron

RF: Random Forests

SOC: State of Charge

SVM: Support Vector Machines

SSE: Sum of Squared Errors

t-SNE: t-Distributed Stochastic Neighbor Embedding

UKF: Unscented Kalman Filter

List of symbols:

CA: Set of charging segments for Car A

n: Total number of charging segments for Car A

C<sub>i</sub>: The i-th charging segment of CA

T<sub>ij</sub>: Sampling time at the j-th point in C<sub>i</sub>, s

I<sub>ij</sub>: Current at time t<sub>ij</sub>, A

$S_{ij}$ : State of charge at time  $t_{ij}$ , %

$T_{ij}$ : Temperature at time  $t_{ij}$ , °C

$V_{ij}$ : Voltage at time  $t_{ij}$ , V

$\Delta t(k)_i$ : Time interval within the  $k$ -th SOC sub-segment of  $C'_i$ , s

$\bar{I}(k)_i$ : Average current within the  $k$ -th SOC sub-segment of  $C'_i$ , A

$\bar{V}(k)_i$ : Average voltage within the  $k$ -th SOC sub-segment of  $C'_i$ , V

$\bar{T}(k)_i$ : Average temperature within the  $k$ -th SOC sub-segment of  $C'_i$ , V

$n(k)_i$ : Number of sampled data points within the  $k$ -th SOC sub-segment of  $C'_i$

GSM: Set of global statistical feature vectors for  $C'_i$

$C'_i$ : Sequence of statistical features extracted from  $C_i$

$\tilde{C}_A$ : Set of all transformed segments for Car A

$K_i$ : Number of SOC sub-segments in  $C'_i$

$X$ : Filled matrix after low-rank matrix completion

$\Omega$ : Indicator matrix for observed entries

$\lambda$ : Regularization parameter

$x(a)_i$ : Original value of the  $a$ -th feature in  $C'_i$

NTSM: Standardized time-series feature matrix for  $C'_i$

$T_i$ : Standardized matrix of all features for  $C'_i$

$T(a)_i$ : Standardized value of the  $a$ -th feature in  $C'_i$

NGSM: Set of all standardized global statistical features for  $C'_i$

$H_i$ : Multi-scale feature set extracted from  $C'_i$

$W_r$ : Convolutional kernel of size  $r \times 1$

$W'$ : Weight matrix

$W''$ : Weight matrix

$F1$ : Deep feature vector extracted from  $C'_i$

$F2$ : Comprehensive feature vector of  $C'_i$  obtained by fusing  $F1$  and NGSM

PC: Set of charging pattern labels predicted for C'i

Cnew: A real-time new charging segment for Car A

C'new: Transformed statistical feature sequence of Cnew

S: Initial prediction curve

Sadjusted: Dynamically adjusted prediction curve

$\delta$ : Set of bias correction terms

$\delta_q$ : Bias correction term at the q-th time step

$\Delta a_q$ : Bias correction term using linear adjustment

$\Delta b_q$ : Bias correction term using logarithmic adjustment

$\Delta c_q$ : Bias correction term using proportional adjustment

$\Delta d_q$ : Bias correction term predicted by the deep learning model

Inewq: Charging current at the q-th moment in the new charging segment

Inewq+1: Charging current at the (q+1)-th moment in the new charging segment

Snewq: SOC at the q-th moment in the new charging segment

snewq+1: SOC at the (q+1)-th moment in the new charging segment

tq: Time at the q-th moment in the new charging segment

## Acknowledgements

This work is supported by the Science and Technology Innovation Key R&D Program of Chongqing (CSTB2023TIAD-STX0024) and the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant number [KJQN202201121](#)).

## Conflict of interest

None

## Financial support

None

## Ethics statement

None

Received: 18 Sep 2021

Revised: 16 Oct 2021

Accepted: 29 Nov 2021

Published online: 20 March 2022

## Rights and permissions

**Open Access** The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

## References

- International Energy Agency. Global EV Outlook 2024. Paris: International Energy Agency; 2024. <https://www.iea.org/reports/global-ev-outlook-2024>
- Frendo O, Graf J, Gaertner N, Stuckenschmidt H. Data-driven smart charging for heterogeneous electric vehicle fleets. *Energy AI*. 2020;1:100007.
- Qu H, Kuang H, Wang Q, Li J, You L. A physics-informed and attention-based graph learning approach for regional electric vehicle charging demand prediction. *IEEE Trans Intell Transp Syst*. 2024;25(10):14284-97.
- Cao X, Hu P, Liu D. Progress of research on electric vehicle charging stations. *Prog Geogr*. 2019;38:139-52.
- Khan AB, Pham VL, Nguyen TT, Choi W. Multistage constant-current charging method for Li-ion batteries. In: 2016 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific). IEEE; 2016. p. 381-5.
- Lavety S, Keshri RK, Chaudhari MA. Multistep constant current-constant voltage charging strategy for a valve regulated lead-acid battery. *IEEE Trans Ind Appl*. 2021;57(6):6494-503.
- Acharige SS, Haque ME, Arif MT, Hosseinzadeh N, Hasan KN, Oo AMT. Review of electric vehicle charging technologies, standards, architectures, and converter configurations. *IEEE Access*. 2023;11:41218-55.
- Suarez C, Martinez W. Fast and ultra-fast charging for battery electric vehicles—a review. In: 2019 IEEE Energy Conversion Congress and Exposition (ECCE). IEEE; 2019. p. 569-75.
- Koli A, Mundotiya P, Tiwari H and Dave V. An improved CCCV controlled charging technique for electric vehicles. In: 2023 International Conference on Computer, Electronics and Electrical Engineering and their Applications (IC2E3). IEEE; 2023. p. 1-5.
- Sabarimuthu M, Senthilnathan N, Kamalesh MS. Multi-stage constant current–constant voltage under constant temperature (MSCC-CV-CT) charging technique for lithium-ion batteries in light weight electric vehicles (EVs). *Electr Eng*. 2023;105(6):4289-309.
- Madaram VG, Biswas PK, Sain C, Thanikanti SB, Balachandran PK. Advancement of electric vehicle technologies, classification of charging methodologies, and optimization strategies for sustainable development - a comprehensive review. *Heliyon*. 2024;10(20).
- Liu H, Naqvi IH, Li F, Liu C, Shafiei N, Li Y, Pecht M. An analytical model for the CC-CV charge of Li-ion batteries with application to degradation analysis. *J Energy Storage*. 2020;29:101342.
- Shi J, Tian M, Han S, Wu TY, Tang Y. Electric vehicle battery remaining charging time estimation considering charging accuracy and charging profile prediction. *J Energy Storage*. 2022;49:104132.

Müller V, Kaiser R, Poller S, Sauerteig D. Importance of the constant voltage charging step during lithium-ion cell formation. *J Energy Storage*. 2018;15:256-65.

Ge X, Shi L, Fu Y, Muyeen SM, Zhang Z, He H. Data-driven spatial-temporal prediction of electric vehicle load profile considering charging behavior. *Electr Power Syst Res*. 2020;187:106469.

Yang C, Zhou H, Chen X, Huang J. Demand time series prediction of stacked long short-term memory electric vehicle charging stations based on fused attention mechanism. *Energies*. 2024;17(9):2041.

Chen Y, Tang Z, Cui Y, Rao W, Li Y. Electric vehicle charging demand prediction model based on spatiotemporal attention mechanism. *Energies*. 2025;18(3):687.

Tian J, Liu H, Gan W, Zhou Y, Wang N, Ma S. Short-term electric vehicle charging load forecasting based on TCN-LSTM network with comprehensive similar day identification. *Appl Energy*. 2025;381:125174.

Kreft M, Brudermueller T, Fleisch E, Staake T. Predictability of electric vehicle charging: explaining extensive user behavior-specific heterogeneity. *Appl Energy*. 2024;370:123544.

Xu X, Wu D, Yang L, Zhang H, Liu G. State estimation of lithium batteries for energy storage based on dual extended Kalman filter. *Math Probl Eng*. 2020;2020:6096834.

Liu J, Guo H, Xiong J, Kato N, Zhang J, Zhang Y. Smart and resilient EV charging in SDN-enhanced vehicular edge computing networks. *IEEE J Sel Areas Commun*. 2019;38(1):217-28.

Shahriar S, Al-Ali AR, Osman AH, Dhou S, Nijim M. Prediction of EV charging behavior using machine learning. *IEEE Access*. 2021;9:111576-86.