World Electric Vehicle Journal

MDPI

Article

A Time-Partitioned Dual-Layer LSTM Based on Route Spatiotemporal for Electric Bus Energy Prediction

Yue Wang [1,2,3], Yu Wang [1,3], Shiqi Liu [2], Yanpeng Zhu [2], Bo Wang [2,*], Yixin Li [1], Guoqun Yao [2] and Wei Zhong [3,*]

[1] School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China; wangyue@ustb.edu.cn (Y.W.); m202510490@xs.ustb.edu.cn (Y.W.); 18836935602@163.com (Y.L.)
[2] Zhongtong Bus Holding Co., Ltd., Liaocheng 252000, China; liusq@zhongtong.com (S.L.); zhuyp@zhongtong.com (Y.Z.); 13963008258@163.com (G.Y.)
[3] State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University, Beijing 100084, China
* Correspondence: wangbo510@zhongtong.com (B.W.); zhongwei@mail.tsinghua.edu.cn (W.Z.)

Abstract

Existing energy consumption models suffer from accuracy degradation and limited robustness in complex urban environments due to insufficient consideration of the route spatiotemporal characteristics of electric buses. To address this limitation, a Time-Partitioned Dual-Layer LSTM (TP-D-LSTM) framework driven by cloud data and spatiotemporal characteristics is proposed. First, a spatiotemporal characteristics analysis is conducted on urban bus routes to reveal the underlying traffic flow dynamics. Based on these insights, a time-partitioning strategy is developed to classify the continuous operating data into independent periods while preserving the kinematic continuity of individual trips. Subsequently, a Dual-Layer LSTM (D-LSTM) is constructed to precisely capture the distinct energy consumption mechanisms within each partitioned scenario. Experiments based on real-world cloud-logged data demonstrate that the proposed TP-D-LSTM framework is superior to existing baseline models. By alleviating the limitations of global mixed modeling, the TP-D-LSTM significantly reduces the Root Mean Square Error (RMSE) to 6.15, achieving an improvement of over 50% compared to the D-LSTM, and exhibits remarkable stability under highly volatile traffic conditions.

Keywords: electric bus; route; energy consumption prediction; dual-layer LSTM

1. Introduction

Electric buses (EBs) have emerged as a cornerstone technology in the global transition toward sustainable public transportation, offering significant advantages in reducing tailpipe emissions and improving overall energy efficiency [1,2]. Driven by stringent decarbonization goals, the large-scale deployment of EBs is rapidly accelerating worldwide [3]. However, the practical operation of EB introduces critical challenges, primarily stemming from range anxiety under complex traffic. Consequently, accurate energy consumption prediction has become a fundamental prerequisite for reliable fleet management. Developing robust forecasting models is not only vital for maintaining service regularity and preventing unexpected battery depletion during operation [4], but also essential for broader integrated energy system optimizations. These include optimizing charging-station planning [5] and enabling intelligent charging–scheduling frameworks to minimize operational costs [6,7]. Therefore, deeply investigating and improving the accuracy of EB energy consumption prediction is of paramount importance to ensure the operational continuity and economic viability of modern transit systems.

Check for updates

Academic Editor: Michael Fowler

Received: 11 March 2026
Revised: 9 April 2026
Accepted: 13 April 2026
Published: 16 April 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the World Electric Vehicle Association. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Achieving high-precision prediction requires explicitly accounting for diverse real-world factors. For examples, variable passenger loads [8] and extreme weather conditions [9] significantly influence overall energy consumption. Furthermore, driving-cycle characteristics—such as speed fluctuations and stop-and-go patterns—strongly shape instantaneous traction power [10]. Fundamentally, the severe volatility in these operational factors is inherently driven by the deeply coupled spatiotemporal heterogeneity of urban transit networks.

From a temporal perspective, the operational efficiency of transit routes is highly susceptible to time-varying traffic dynamics. Specifically, peak-hour congestion forces EBs into prolonged low-speed operation, frequent acceleration–deceleration cycles, and extended idling, drastically increasing energy intensity per kilometer [11]. Furthermore, morning and evening rush-hour traffic waves induce unstable flow conditions, directly amplifying the temporal heterogeneity of EB energy profiles [12]. During these periods, fluctuating queueing times and stop delays introduce severe variations in instantaneous traction power demand [13].

From a spatial perspective, EB energy profiles are further complicated by the topological heterogeneity of the road network. Specifically, physical route features—including intersection geometry, signalized corridors, and varying bus-stop spacing—create highly localized peaks in power demand [14]. Furthermore, road-grade variations drastically amplify these spatial discrepancies, as uphill segments demand substantial propulsion power while downhill slopes dictate regenerative-braking efficiency [15].

As traditional models struggle to capture deeply coupled spatiotemporal dynamics, deep learning has been widely adopted for bus energy prediction. Existing methodologies can be categorized into two primary dimensions: spatiotemporal data modeling and internal mapping mechanisms.

First, representative deep learning approaches for bus energy prediction mainly focus on processing continuous all-day operational data. For instance, LSTM-based models have been applied to predict energy consumption by learning holistic temporal patterns from historical driving data [16]. CNN-based frameworks have been explored for estimation leveraging time-series feature extraction across the daily cycle [17]. Moreover, advanced architectures, including foundation models, have been recently introduced to further enhance spatiotemporal sequence modeling capabilities for complex intelligent transportation tasks [18]. Similarly, other recent deep learning frameworks have expanded these sequence modeling capabilities to capture global operational trends over entire routes [19,20]. Thus, it can be seen that these representative methods predominantly employ a unified global model to extract average temporal features across the entire daily cycle, which inevitably leads to a severe "pattern averaging" effect when handling highly heterogeneous traffic conditions such as peak-hour congestion.

Second, current deep learning approaches rely on pure data-driven end-to-end mapping. For instance, recent studies and surveys highlight that the energy consumption process is frequently formulated as a pure data-driven mapping from external features to power output [21]. Similarly, to enhance prediction performance, various advanced network architectures are widely adopted to construct implicit energy relationships based on large-scale operational data [22,23]. Moreover, recent frameworks have employed pre-trained big data models and complex stacking strategies to continuously minimize fitting errors, focusing on data correlations [24,25]. Although these data-driven strategies significantly improve prediction accuracy through complex neural networks, their inherent "black-box" nature makes it difficult to explicitly reflect the underlying electrochemical mechanisms, thereby limiting the physical interpretability of the models.

To overcome these methodological bottlenecks, acquiring comprehensive, high-granularity operational data is a fundamental prerequisite. Fortunately, emerging technologies offer new avenues for addressing these issues. Vehicle–cloud collaboration has been shown to enable joint state-of-energy estimation by leveraging cloud-logged battery history and operational data to improve predictive accuracy [26]. A cloud energy management strategy incorporating edge-cloud collaboration architectures has also been proposed for intelligent connected vehicles, demonstrating the utility of hierarchical cloud platforms for energy optimization [27]. Moreover, recent big data-empowered frameworks exploit heterogeneous EV data via edge-cloud collaboration to support scalable analytics and decision-making [28]. Based on this framework, this paper proposes a novel Time-Partitioned Dual-Layer LSTM (TP-D-LSTM) framework based on cloud-logged operational data. The primary contributions of this paper can be summarized as follows:

(1) A cloud-based data acquisition and processing framework is proposed to handle the distinct characteristics of cloud-side data. Unlike vehicle-end data acquisition, this approach accounts for the heterogeneity of multi-source information in the cloud-platform, ensuring high-quality inputs for prediction.
(2) A spatiotemporal-characteristics-based route feature extraction is developed to decouple the inherent heterogeneity of urban bus operations. Compared with conventional route analysis, it preserves temporal continuity across regimes and mitigates the "pattern averaging" effect.
(3) A Time-partitioned dual-layer LSTM (TP-D-LSTM) architecture is developed for high-precision prediction. Different from standard deep learning models, it incorporates intermediate electrochemical state prediction and residual correction, improving both interpretability and predictive accuracy under complex traffic conditions.

This paper is organized as follows. Section 2 introduces the bus route data acquisition and preprocessing. Section 3 presents the analysis of spatiotemporal characteristics. Section 4 details the proposed time-partitioned dual-layer LSTM. Section 5 discusses the experimental results. Finally, Section 6 concludes the paper.

2. Acquisition and Preprocessing of Cloud-Based Operating Data

To establish a reliable empirical foundation for energy consumption modeling, this section systematically describes the data acquisition and processing framework. First, the architecture of the vehicle–road–cloud collaborative platform and the technical specifications of the experimental object (Jinan K260 electric bus) are introduced. Subsequently, the multi-dimensional dataset fusing GPS trajectory and CAN bus parameters is defined. Finally, a preprocessing pipeline is detailed, with a specific focus on the dwell-time-based trajectory splitting mechanism employed to extract effective operational cycles from continuous raw streams.

2.1. Vehicle–Road–Cloud Platform Architecture and Physical Configurations

As illustrated in Figure 1, the data acquisition and analysis framework is built upon a vehicle–cloud collaborative platform. The vehicle side functions as an edge node to capture real-time operational states and upload them to the server. On the cloud side, road–cloud perception fusion is performed to integrate these vehicle data streams with holographic traffic information collected from roadside infrastructure.

To ensure the representativeness of the experimental data, an electric bus from the Jinan K260 line was selected as the research object. The specific technical specifications of the vehicle, including total mass, battery capacity, and peak motor power, are detailed in Table 1. These parameters provide the physical basis for the subsequent energy modeling.

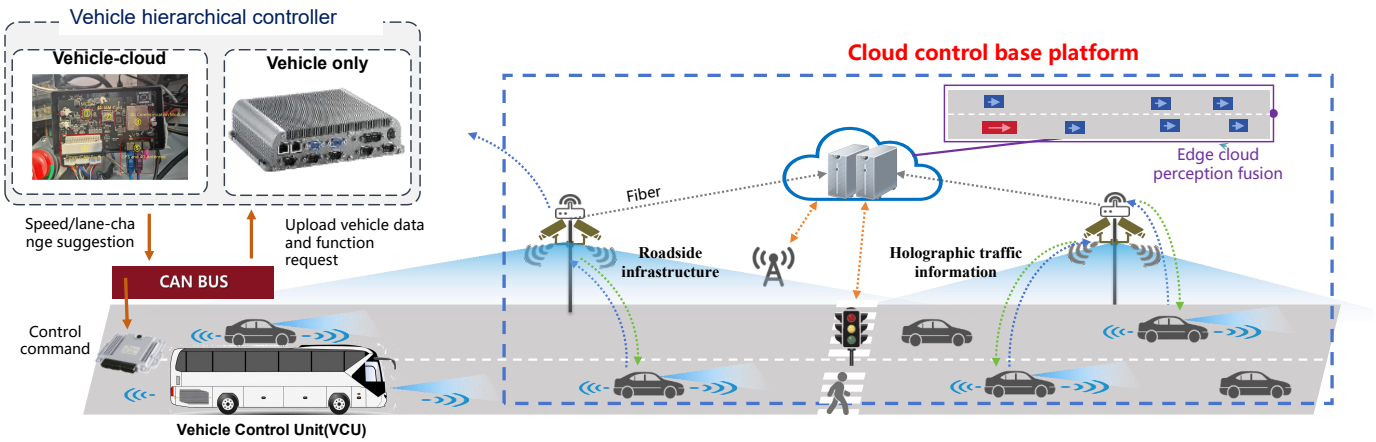


Figure 1. Framework of the vehicle–cloud collaborative system.

Table 1. Technical specifications and feature dimensions of the experimental bus.

Category	Parameter	Symbol	Value/Unit
Static Specs	Overall Dimensions ($L \times W \times H$)	-	$10.5 \times 2.55 \times 3.27$ m
	Frontal Area	A	7.9 m^2
Dynamics	Curb Weight/Gross Vehicle Weight	m/M	12,000/17,200 kg
	Rolling Radius	r	0.446 m
	Final Drive Ratio	i_0	5.67
Dynamic Features	Instantaneous Speed/Acceleration	v/a	km/h; m/s^2
	Longitude and Latitude and Altitude	Lon, Lat, Alt	°, °, m
Power States	Motor Speed/Motor Torque	N/T	rpm; Nm
	Bus Voltage/Bus Current	U/I	V; A
	State of Charge	SOC	%

To further characterize the energy conversion efficiency, the efficiency MAP of the traction motor is integrated into our analysis (see Figure 2). In the actual operation process, the motor speed (RPM) and torque (Tq) captured by the vehicle-mounted terminal realize nonlinear mapping through this efficiency matrix. This mapping relationship can capture the energy loss characteristics of the power system under different load conditions, thus transforming the original current and voltage records into mechanical work output with clear physical significance.

2.2. Dynamic–Static Data Association and Sequence Preprocessing

To comprehensively characterize the energy consumption mechanism, diverse data types are integrated into a unified structure. As shown in Figure 3, the dataset consists of two primary streams: the dynamic trajectory data (L) and the static business data (B). The trajectory data stream is generated by the GPS and CAN bus systems at a frequency of 1 Hz. A continuous sequence is defined as $L = \{L_1, L_2, \ldots, L_n\}$, where each data point L_i contains a multidimensional tuple of attributes:

$$_{+}L_i = \{Vin_i, Time_i, Lon_i, Lat_i, Alt_i, v_{gps,i}, h_i, RPM_i, Tq_i, P_{ins,i}, U_i, I_i, SOC_i, v_{veh,i}\} \quad (1)$$

Specifically, these variables are categorized as follows: Vin_i denotes the unique vehicle identification number, and $Time_i$ is the upload timestamp. The geographical information includes longitude (Lon_i), latitude (Lat_i), and altitude (Alt_i). Kinematic states are captured by the GPS-measured speed ($v_{gps,i}$), vehicle heading (h_i), and the dashboard-displayed

speed ($v_{veh,i}$). Regarding the powertrain, RPM_i and Tq_i represent the motor speed and torque, respectively, while $P_{ins,i}$, U_i, I_i, and SOC_i correspond to the instantaneous power, bus voltage, current, and battery State of Charge.

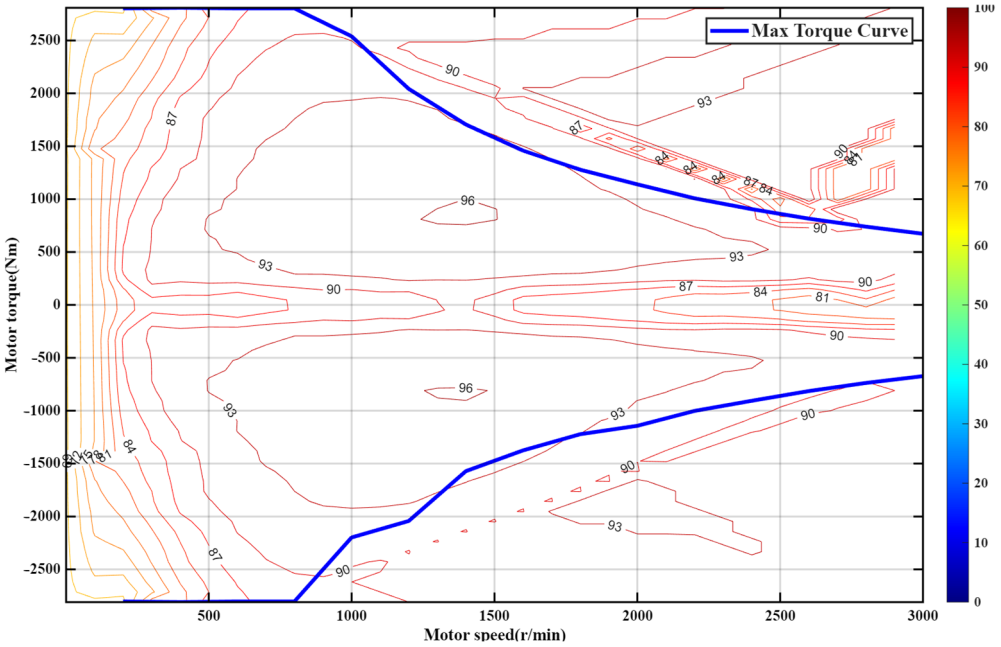


Figure 2. Motor MAP diagram.

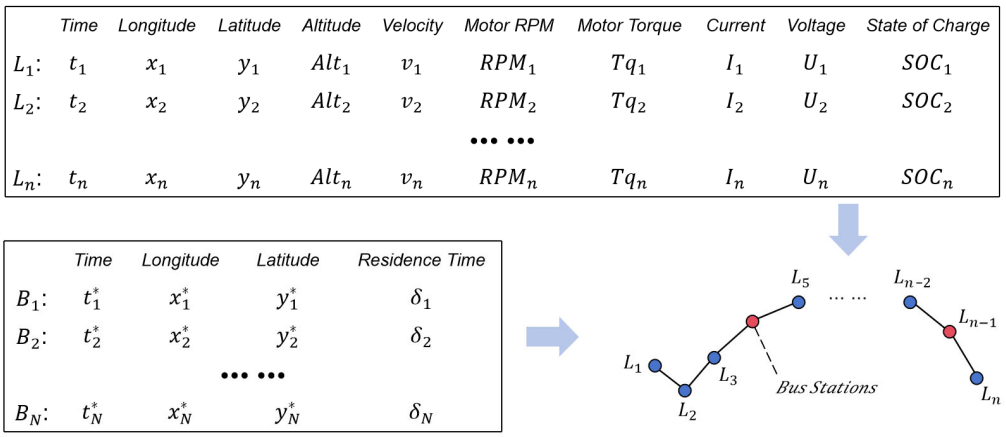


Figure 3. GPS log and trajectory of the bus.

The business data stream, defined as $B = \{B_1, B_2, \ldots, B_N\}$, records the critical moments of vehicle arrival at stations. As detailed in the station table (bottom-left of Figure 3), each record is parameterized as $B_j = \left\{t_j^*, x_j^*, y_j^*, \delta_j\right\}$. Here, t_j^* represents the specific time when the vehicle arrives at the business point, x_j^* and y_j^* refer to the longitude and latitude coordinates of the station, and δ_j denotes the residence time of the vehicle at that point. By spatially correlating the dynamic sequence L with these static business points, a complete data foundation is established.

Raw data transmitted from the vehicle inevitably contains noise, GPS drift, and packet loss due to signal instability. To address these issues, a systematic preprocessing framework is implemented, the overall procedure of which is depicted in Figure 4.

First, a rule-based outlier repair is applied. Velocity spikes (e.g., sudden zero values during motion) are smoothed using a moving average of adjacent points, while unrealistic location jumps that violate kinematic constraints are eliminated to ensure trajectory continuity.

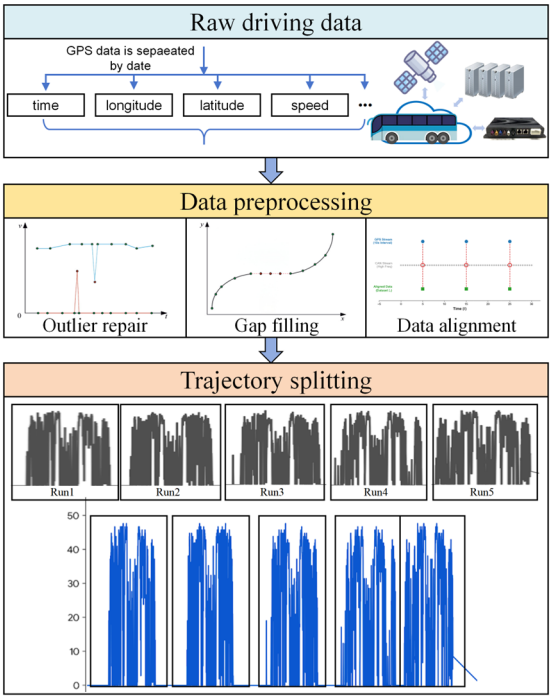


Figure 4. Framework for GPS trajectory data segmentation.

Following basic cleaning, the core task is Trajectory Splitting, which aims to isolate effective operational "Runs" from the continuous stream. The presence of "Idle Points" (records during terminal layovers) can significantly distort travel time analysis. Therefore, instead of complex recursive calculations, we employ a dwell-time-based segmentation strategy. A valid run is defined as a continuous moving sequence separated by a stationary duration exceeding a specific threshold (e.g., terminal stops).

Table 2 demonstrates this splitting mechanism with real-world data. The vehicle arrives at the station at 9 April 2025 10:42:42, after which the velocity drops to zero. The system detects a continuous non-movement interval of 2120 s (labeled as Idle Points). By recognizing this long-duration stop and removing the corresponding idle records (rows 5–10 in Table 2), the continuous sequence is effectively cut into two independent trips.

Table 2. GPS trajectory data.

t	x	y	RPM	v
9 April 2025 10:42:02	116.8950757	36.81136350	1252	37.82
9 April 2025 10:42:12	116.8950202	36.81193800	472	13.66
9 April 2025 10:42:22	116.8946093	36.81196167	841	25.21
9 April 2025 10:42:32	116.8941537	36.81164850	745	22.06
9 April 2025 10:42:42	116.8941325	36.81140883	0	0
9 April 2025 10:50:22	116.8939383	36.81119217	0	0
9 April 2025 10:56:12	116.8939312	36.81119550	0	0
9 April 2025 11:03:42	116.8939390	36.81118450	0	0
9 April 2025 11:09:02	116.8939362	36.81117967	0	0
9 April 2025 11:17:52	116.8939383	36.81122583	0	0
9 April 2025 11:18:02	116.8940685	36.81131333	404	11.55
9 April 2025 11:18:12	116.8945442	36.81118950	404	11.55
9 April 2025 11:18:22	116.8948353	36.81112450	451	13.66
9 April 2025 11:18:32	116.8949415	36.81095317	480	13.66

3. Analysis of Spatiotemporal Characteristics for Urban Bus Routes

To construct a high-precision energy prediction model, it is essential to quantify the operational characteristics of the bus route. This section establishes a framework for extracting features from static physical attributes to dynamic spatiotemporal patterns, taking the Jinan K260 line as the research object.

3.1. Route Profiling and Microscopic Kinematic Indicators

The K260 bus line represents a typical suburban–urban transition route, spanning approximately 20 km from Sangzidian Bus Depot to the Zoo. Figure 5 illustrates the spatial distribution of the 22 stations and the count of traffic lights along the route. Complementing this, Figure 6 details the inter-station distances and key station attributes. Based on these physical constraints, the route is divided into two distinct sections:

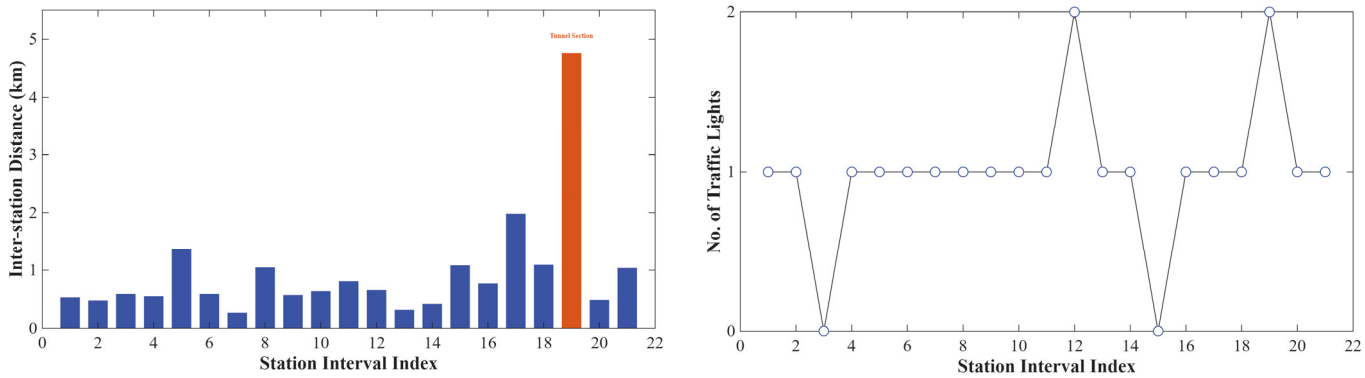


Figure 5. Spatial distribution of Route K260: (**a**) inter-station distances; (**b**) traffic light counts (markers indicate data points).

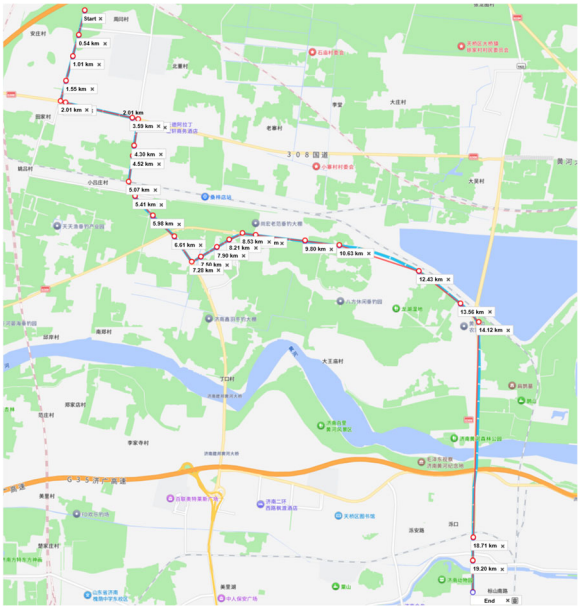


Figure 6. Real-world Statistics of inter-station distances and station attributes. (Note: Background Chinese characters are merely local POIs, not affecting the route's scientific meaning).

Ground Section: Characterized by high station density and frequent signalized intersections (as shown in the high red-light count in Figure 5), leading to frequent stop-and-go maneuvers.

Tunnel Section: Corresponds to the Jiluo Road Yellow River Tunnel (approx. 13.3–18.1 km). This segment features a closed environment with zero traffic lights and long inter-station distances (see Figure 6), allowing for sustained high-speed cruising.

To quantify the operational status, we extract kinematic fragments from the raw trajectory. Figure 7 compares the typical velocity waveforms of different sections. The ground section exhibits fragmented, short-duration waves due to signal interruptions, whereas the tunnel section presents long-period trapezoidal cruising waves.

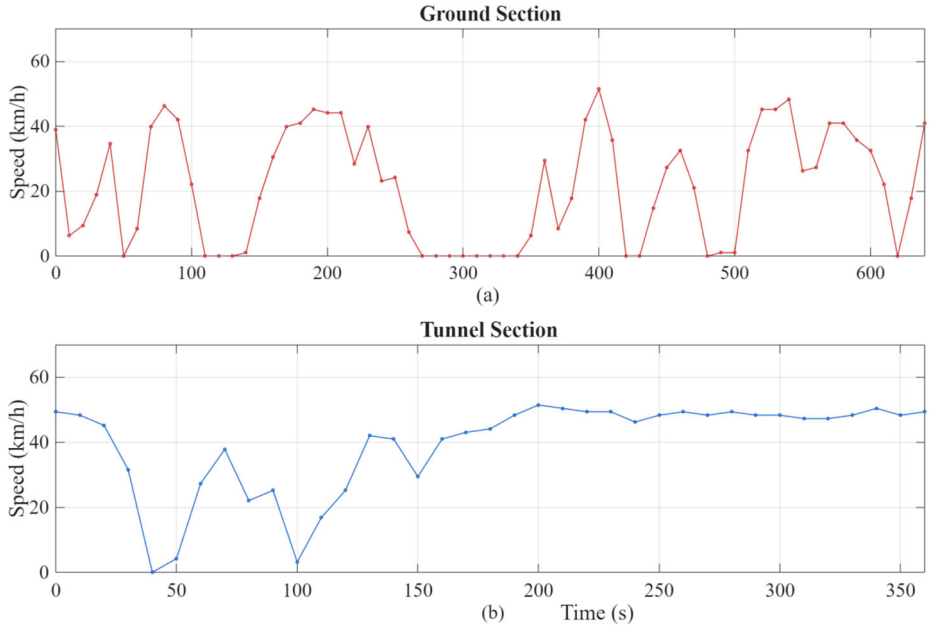


Figure 7. Comparison of velocity profiles: (**a**) ground section; (**b**) tunnel section.

Based on the extracted kinematic fragments, three key microscopic indicators are defined to quantitatively characterize the traffic conditions. First, the low-speed ratio is introduced to measure congestion intensity. It represents the proportion of operational time where the vehicle speed falls below 25 km/h, physically corresponding to high-energy-consumption regimes such as creeping and queuing. Second, the stop density, defined as the number of stops per kilometer, is utilized to reflect the spatial intensity of physical blockages caused by stations and signals. Finally, velocity volatility, calculated as the standard deviation of the speed sequence, serves as a comprehensive indicator to characterize the non-stationarity and stability of the traffic flow.

3.2. Spatiotemporal Coupling and Long-Tail Effect Analysis

Constrained by both the tidal effects of commuting and the physical road environment, the operational status of urban bus routes exhibits significant non-stationary and stochastic characteristics. To reveal the underlying mechanisms of energy consumption fluctuations in the K260 route, this section utilizes high-frequency trajectory data to analyze the dynamic features of traffic flow from two dimensions—temporal evolution and spatial distribution—and demonstrates the impact of spatiotemporal coupling effects on vehicle energy consumption.

3.2.1. Analysis of Temporal Evolution Characteristics

The temporal variation in operating speed reflects the dynamic fluctuations of macroscopic traffic flow density. Figure 8 illustrates the time-varying characteristics of the average driving speed and velocity volatility for the K260 route over the entire month. Statistical results indicate that the daily operation exhibits significant non-stationary stochastic features,

with a strong negative correlation between the mean velocity and volatility. Specifically, volatility decreases significantly when the average speed maintains a high level above 30 km/h, whereas it surges when the speed drops into a trough.

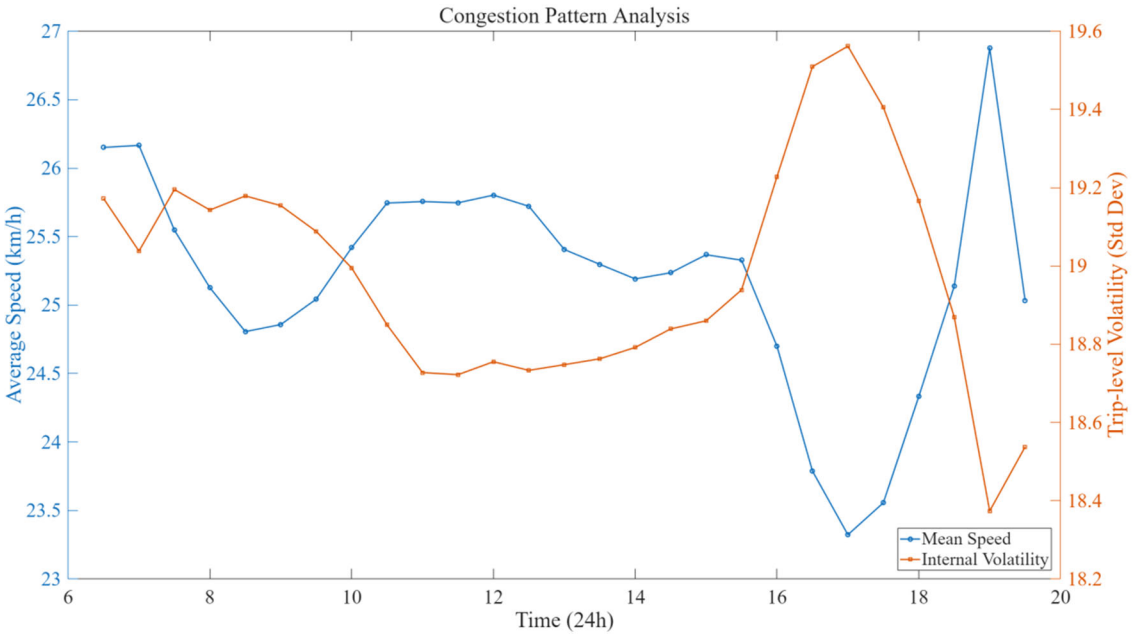


Figure 8. Temporal evolution of average speed and velocity volatility.

Notably, the velocity profiles of morning and evening peaks exhibit significant asymmetry. While the Morning Peak displays a "steep and concentrated" pattern with rapid dissipation, the Evening Peak presents a distinct "deep valley and long-tail" morphology. This "long-tail effect" indicates that the traffic demand on ground sections exceeds the saturation threshold of road capacity, causing the queue dissipation rate to lag significantly behind the aggregation rate. Consequently, the evening peak represents the most severe non-linear operating scenario, characterized by cumulative congestion.

To objectively define the typical operating periods, a data-driven approach is adopted based on the monthly aggregated velocity profile (as shown in Figure 8). Specifically, the critical threshold of 25 km/h—a standard indicator of traffic congestion—is utilized to segment the daily operation cycle.

As observed in Figure 8, the curve exhibits two distinct troughs where the average speed falls below this threshold. Based on these inflection points, the operating hours are categorized as follows:

Morning Peak (08:12–09:23): Corresponds to the first congestion trough, spanning approximately 1.1 h.

Evening Peak (15:46–18:25): Represents the deepest and widest trough with a duration of 2.6 h. This period exhibits the most significant long-term dependencies and volatility.

Off-peak Period: The remaining intervals where the vehicle speed rebounds significantly above 25 km/h, indicating a stabilized traffic flow.

Table 3 further quantifies the microscopic operational characteristics of these phases. Statistical data indicate that the Morning Peak exhibits the highest traffic intensity, with the Stop Density reaching 1.81 stops/km and the Low-Speed Ratio peaking at 67.83%. In comparison, while the Evening Peak shows slightly lower average intensity metrics (1.38 stops/km and 63.77%), it sustains these high-congestion conditions for a much longer duration (2.6 h vs. 1.1 h). This distinct characteristic—highest intensity in the morning versus longest saturated duration in the evening—confirms the non-uniformity of daily

traffic and provides direct data support for the necessity of constructing a time-partitioned prediction framework.

Table 3. Microscopic characteristics of three typical operating periods.

Period	Avg Speed (km/h)	Low Speed Ratio (%)	Stop Density (Stops/km)
Morning Peak	21.99	67.83	1.81
Off-Peak	25.64	57.49	1.09
Evening Peak	22.99	63.77	1.38

3.2.2. Analysis of Spatial Distribution Characteristics

Single-dimensional temporal analysis is insufficient to fully characterize the complex operating conditions, as congestion is often the result of the coupling between specific time windows and specific spatial nodes. To investigate this spatiotemporal coupling mechanism, we constructed spatiotemporal velocity heatmaps based on the full-month operational data. As illustrated in Figure 9, the horizontal axis represents the driving mileage, the vertical axis represents the date, and the color scale indicates the instantaneous speed.

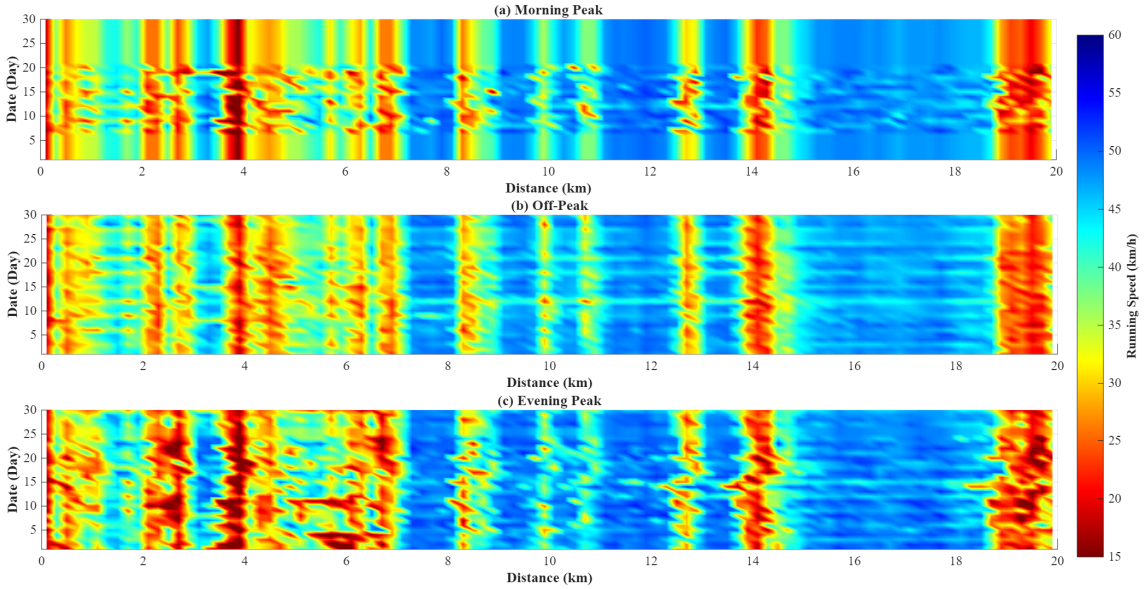


Figure 9. Spatiotemporal velocity heatmaps of full-month operations: (**a**) Morning Peak period, (**b**) Off-Peak period, and (**c**) Evening Peak period.

By comparing the heatmaps across different periods, distinct spatiotemporal evolution patterns are observed. First, the "Spatial Fixity of Bottlenecks" is evident. In all three subplots, vertical color stripes persist across the entire timeline. Specifically, the 0–5 km starting section exhibits continuous red/orange low-speed bands. This confirms the existence of fixed physical bottlenecks, such as high-density signalized intersections, whose congestion attributes are spatially constrained and independent of the date.

In sharp contrast, the "Temporal Invariance of the Tunnel Section" is observed in the 8–12 km and 14–18 km segments. These areas display stable deep blue bands with speeds consistently above 40 km/h across all periods.

Furthermore, comparing Figures 9a and 9c reveals the "Extreme Conditions of the Evening Peak." Although the morning and evening peaks share the same spatial bottlenecks, the congestion intensity in the evening is significantly higher. Figure 9c displays extensive deep red to black-red patches representing speeds below 15 km/h in the

0–6 km region, with a wider lateral span compared to the morning peak. This visual evidence corroborates the statistical conclusion in Section 3.2.1 regarding the "cumulative congestion effect".

In summary, the energy consumption peaks of the K260 line are generated by the deep coupling of the "Evening Peak time window" and "Ground Section spatial bottlenecks." This binary spatial structure—where local extreme congestion on the ground and local free flow in the tunnel coexist—results in highly heterogeneous data distributions. Given that the physical route topology is fixed, the operational variability is primarily driven by temporal dynamics. Consequently, a decoupling strategy based on temporal segmentation is sufficient to effectively isolate these distinct flow patterns (e.g., separating the volatile congestion on ground sections during evening peaks), thereby mitigating the "pattern averaging" problem inherent in global models.

4. Time-Partitioned Dual-Layer LSTM Framework

Based on the obvious working conditions difference between morning and evening peak periods and flat peak periods revealed in Section 3. This study proposes a "Time-Partitioned Dual-Layer LSTM Framework." The framework operates in two stages: first, a rule-based soft-partitioning strategy decouples continuous data into independent scenarios; second, a dual-layer LSTM architecture is constructed to capture the distinct non-linear energy mapping mechanisms within each scenario.

4.1. Trip-Based Soft Time-Partitioning Strategy

Traditional time-based partitioning often employs "hard thresholds" (e.g., strictly cutting data at 09:23). However, for LSTM models that rely on long-term temporal dependencies, such hard cuts can abruptly interrupt ongoing trips, destroying the kinematic continuity and the causal chain of vehicle motion.

To achieve scenario decoupling while preserving the temporal integrity required by deep learning models, we propose a soft-partitioning strategy based on overlap ratio. Instead of slicing by independent timestamps, this strategy treats a "complete trip" as the fundamental unit. Specifically, for each trip, we calculate the proportion of its duration that overlaps with the standard peak windows defined in Section 3.2.1. If the overlap ratio exceeds 40%, the trip is deemed to be significantly affected by peak congestion and is classified into the corresponding Peak Dataset; otherwise, it is classified into the Off-Peak Dataset. This approach ensures that every trajectory fed into the LSTM contains a complete start-stop cycle and the evolution of driving behavior, avoiding model memory failure caused by data fragmentation.

Based on this strategy, two categories of datasets are constructed for experimental validation: the Global Mixed Dataset, which contains all raw trips for baseline comparison, and the Spatiotemporal Decoupled Datasets, comprising three independent subsets (Morning Peak, Evening Peak, and Off-Peak) to facilitate the fine-grained modeling in the subsequent section.

4.2. TP-D-LSTM Architecture with Soft Physical Constraints

To enhance prediction accuracy under complex urban conditions, a Dual-Layer LSTM (TP-D-LSTM) architecture is designed. The core motivation is to integrate physical consistency into data-driven modeling. Direct power prediction often treats the powertrain as a black box, ignoring the fact that busbar voltage (U) and battery current (I) are the critical physical intermediates governing power consumption $P = U \cdot I$. Therefore, as illustrated in Figure 10, the proposed model adopts a hierarchical structure. The first layer serves as the Electrochemical State Prediction Layer, explicitly predicting the intermediate physical

states based on kinematic inputs. The second layer functions as the Energy Consumption Refinement Layer, which takes the predicted physical states combined with original kinematic features to output the final energy consumption.

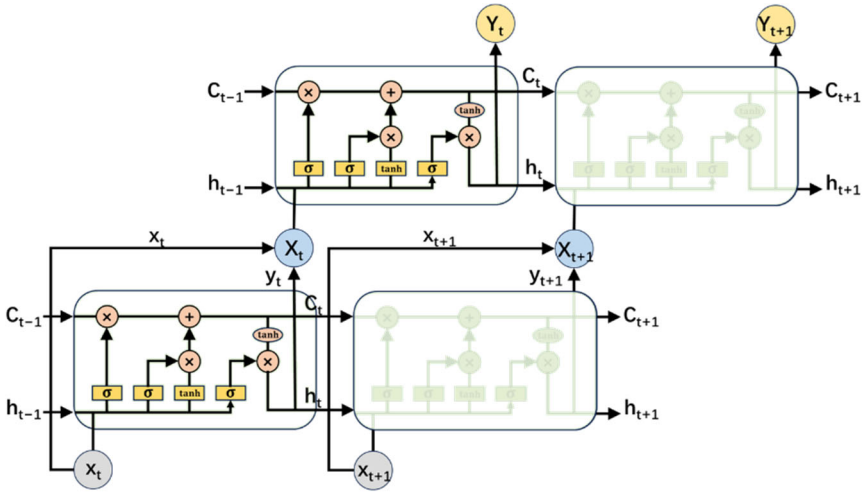


Figure 10. Architecture of the proposed Time-Partitioned Dual-Layer LSTM model.

The fundamental internal gate mechanisms of LSTM units have been well established in the literature [12,29,30], a brief mathematical formulation is presented here to define the underlying temporal engine of our proposed dual-layer architecture. For a given time step t, the information flow—regulated by the forget gate f_t, input gate i_t, and output gate o_t—is governed by the following recursive equations:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \odot \tanh(C_t) \tag{7}$$

where σ denotes the logistic sigmoid function, tanh is the hyperbolic tangent function, and $\odot$ represents the element-wise multiplication. The terms W and b correspond to the learnable weight matrices and bias vectors, respectively.

Within the proposed TP-D-LSTM framework, the input vector x_t carries distinct physical meanings across the two layers: in Layer 1, x_t^1 consists of the time-partitioned kinematic sequences (e.g., speed and acceleration) to predict intermediate electrochemical states; in Layer 2, x_t^2 incorporates these intermediate states along with theoretical power formulas to perform residual correction, ultimately generating the final energy consumption sequence.

The First Layer (Physical State Layer) maps the external kinematic and powertrain states to the internal electrochemical response of the battery. The input vector x_t^1 at time step t includes vehicle velocity (v_t), acceleration (a_t), road slope (θ_t), SOC (SOC_t), motor speed (n_t), and torque (T_t). The output vector $\hat{y}_t^1$ predicts the current and voltage:

$$x_t^1 = [v_t, a_t, \theta_t, SOC_t, n_t, T_t]^T, \hat{y}_t^1 = [\hat{I}, \hat{U}]^T \tag{8}$$

The Second Layer (Energy Consumption Layer) integrates the predictions from the first layer. Crucially, we introduce a physics-guided term $\widetilde{P}_t$, which is the theoretical power calculated from the first layer's output $\widetilde{P}_t = (\hat{I} \cdot \hat{U})/1000$. The input vector x_t^2 is constructed by concatenating the original features, the predicted physical states, and this theoretical power:

$$x_t^2 = [x_t^1, \hat{y}_t^1, \widetilde{P}_t]^T \tag{9}$$

The inclusion of $\widetilde{P}_t$ brings the total input dimension to 9, providing a strong prior knowledge base for the network. The final output $\hat{y}_t^2$ represents the actual instantaneous power $\hat{P}_t$, where the LSTM network learns to correct the residual errors between the theoretical calculation and the actual measurement:

$$\hat{y}_t^2 = [\hat{P}_t] \tag{10}$$

By explicitly incorporating the intermediate variables and the physical formula constraint, the TP-D-LSTM framework imposes a "soft physical constraint" on the data-driven regression, thereby improving both interpretability and robustness against data noise.

5. Experimental Validation and Discussion

This section evaluates the prediction performance and generalization ability of the proposed TP-D-LSTM using real-world bus operation data. The evaluation metrics and network configurations are first defined, followed by a quantitative assessment focusing on three dimensions: the effectiveness of the physical constraint mechanism, the necessity of the TP-D-LSTM, and a comparative analysis against baseline models.

5.1. Evaluation Metrics and Experimental Settings

The model errors are measured to evaluate the performance of the forecasts. The errors are calculated at the same scale as the data. Three evaluation metrics were used in the experiment, including root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) to measure the performance of the models. These evaluation metrics are shown in Equations (11)–(13).

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{11}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(\overline{y} - y_i)^2} \tag{13}$$

where N is the total number of samples, $\hat{y}_i$ denotes the predicted value, y_i is the actual observed value, and $\overline{y}$ represents the mean of the observed values.

A differentiated parameter configuration is adopted, as detailed in Table 4. The first layer, designed to prioritize the extraction of complex state features, is configured with 128 hidden units. In contrast, the second layer focuses on residual correction based on strong physical priors; therefore, it employs a streamlined structure of 64 units. The network incorporates a Dropout rate of 0.5 and an L2 regularization coefficient of 0.01.

Table 4. Model parameter settings.

Model Parameter	Layer 1	Layer 2
Input Dimension	6	9
Time Window Size	200	200
Stride Step	10	10
Hidden Units	128	64
Activation	ReLU	ReLU
Dropout Rate	-	0.5
L2 Regularization	-	0.01
Optimizer	Adam	Adam
Initial Learning Rate	0.0005	0.005
Mini-Batch Size	32	32
Max Epochs	200	200
LR Decay Strategy	Piecewise (Period = 100)	Piecewise (Period = 40)

5.2. Performance Evaluation of the TP-D-LSTM Framework

To comprehensively assess the proposed method, we first validate the necessity of the physical constraint mechanism, then explore the advantages of the TP-D-LSTM, and finally analyze the physical adaptability under different traffic flow states.

5.2.1. Mechanism-Oriented Verification of the Dual-Layer Architecture

Figure 11 validates the stepwise prediction performance of the proposed dual-layer architecture. We first analyze the tracking capability of the intermediate electrochemical states in Layer 1. As illustrated in Figure 11a,b, the predicted battery current (Ib) and voltage (Ub) closely track the ground truth measured by the CAN bus. Even during aggressive acceleration phases where current spikes exceed 150 A, Layer 1 maintains high tracking fidelity, providing a reliable electrical foundation for subsequent energy estimation.



Figure 11. Mechanism-oriented verification of the dual-layer model: (**a**) predicted battery current (Ib) in Layer 1; (**b**) predicted battery voltage (Ub) in Layer 1; (**c**) calculated power (P) in Layer 1; (**d**) predicted power (P) in Layer 2.

However, precise electrical tracking does not inherently guarantee energy accuracy due to unmodeled system losses. As shown in Figure 11c, the energy consumption calculated solely via the physical formula $\left(\widetilde{P}\right)$, based on Layer 1 outputs, exhibits systematic bias.

The error analysis reveals that this pure physical approach yields an RMSE of 18.83 and an MAE of 12.56, primarily due to the neglect of dynamic mechanical transmission efficiencies and nonlinear thermal effects under high-load conditions.

To address these deviations, Layer 2 introduces a residual correction mechanism. By learning the hidden patterns between kinematic states and physical errors via hybrid hidden vectors, Layer 2 effectively compensates for the limitations of the simplified formulas. As demonstrated in Figure 11d, the final corrected output ($\hat{P}$) shows superior alignment with the measured data. Quantitatively, the dual-layer architecture reduces the RMSE by 26.6% (from 18.83 to 13.83) and the MAE by 9.6%, with the coefficient of determination (R^2) stabilizing around 0.965. This step-wise improvement confirms that while Layer 1 establishes the physical mapping, Layer 2 is crucial for internalizing complex operational dynamics and ensuring robust prediction under volatile real-world conditions.

5.2.2. Superiority of the Time-Partitioning Strategy

The experimental results (as shown in Table 5 and Figure 11) reveal the inherent limitations of global unified modeling. When a single model is trained using the mixed full-month data, the average RMSE on the test set remains high at 13.83. Conversely, we now emphasize how the soft time-partitioning strategy decouples heterogeneous operational regimes, transforming complex, mixed traffic patterns into simpler sub-distributions. As a result, as shown in Table 6, RMSE drops from 13.83 (global model) to 8.37, 6.15, and 5.85 for Morning Peak, Evening Peak, and Off-Peak periods, respectively. Notably, the model maintains high predictive fidelity even during the most challenging Evening Peak, effectively internalizing the local dynamic patterns induced by stochastic stop-and-go waves. Compared to the global model, this represents a significant reduction of 55.5% and 57.7% (for the evening and off-peak periods). This result confirms that by explicitly partitioning the complex mixed distribution into multiple relatively simple sub-distributions, the decoupling strategy enables the deep learning model to focus on learning the specific energy consumption laws governing each spatiotemporal domain.

Table 5. Error comparison between physical formula and D-LSTM.

Method	RMSE	MAE	R^2
Physics Formula ($\widetilde{P}$)	18.83	9.32	0.895
D-LSTM ($\hat{P}$)	13.83	6.11	0.965
Improvement	+26.6%	+9.6%	-

Table 6. Error evaluation of TP-D-LSTM across different periods.

Period	RMSE	MAE	R^2
All Scenarios (Global)	13.83	6.11	0.965
Morning Peak	8.37	6.24	0.968
Off-Peak	5.85	5.73	0.984
Evening Peak	6.15	5.89	0.981

Based on the high-precision prediction results of the decoupled models, we further analyze the adaptability characteristics under different typical traffic scenarios. As illustrated in Figure 12b, the model exhibits the best prediction performance in the Off-Peak period (RMSE = 5.85, R^2 = 0.984). During this period, while the bus still experiences significant power fluctuations due to scheduled stops and signalized intersections, the absence of severe congestion makes these acceleration and deceleration cycles more regular and predictable. The TP-D-LSTM model accurately captures these inherent route characteristics,

with the predicted curve almost perfectly overlapping the ground truth, demonstrating excellent tracking of the periodic power peaks without overfitting sensor noise.

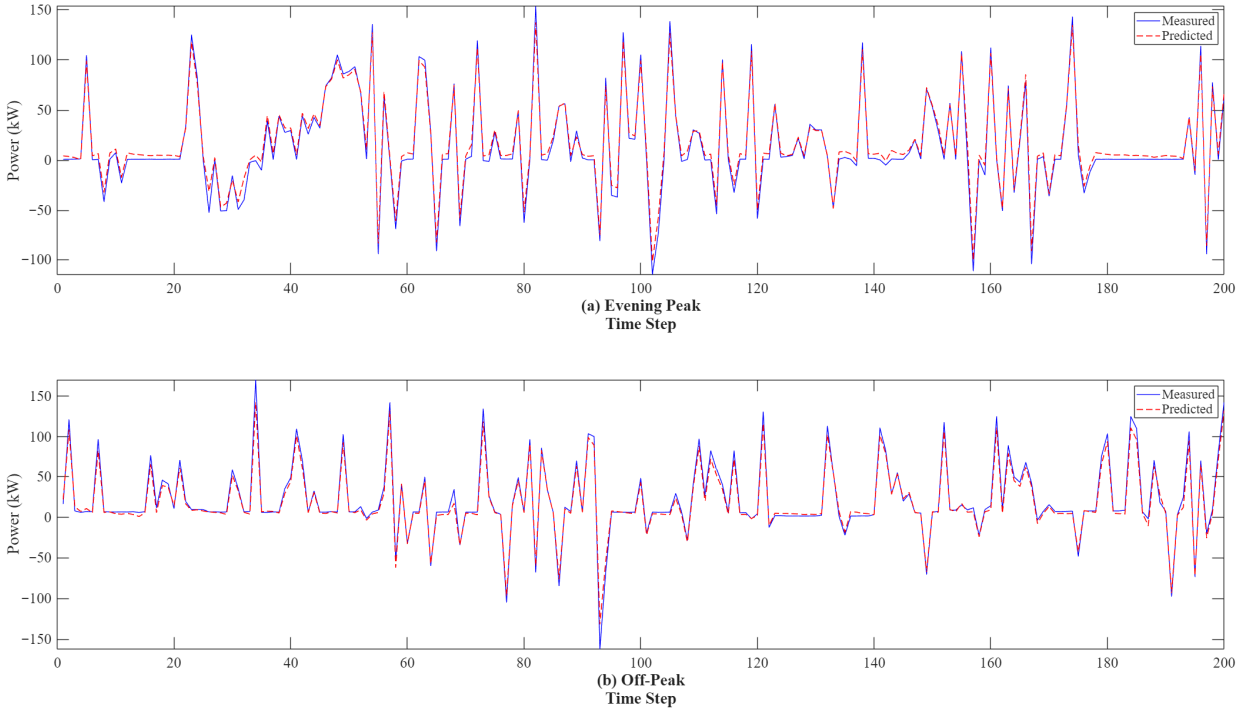


Figure 12. Comparison of prediction performance between different scenarios: (**a**) evening peak; (**b**) off-peak.

In contrast, the Evening Peak period shown in Figure 12a presents a much more challenging scenario. The traffic flow is subject to frequent stop-and-go waves, visually reflected in the denser, high-frequency power jumps and sharper negative spikes. Despite these severe non-linear fluctuations, the model maintains extremely high robustness (RMSE = 6.15). As observed in the figure, even when facing these chaotic transient spikes and fast-changing edges induced by traffic saturation, the model remains highly sensitive, proving that it has effectively internalized the dynamic powertrain response laws under extreme traffic volatility.

5.2.3. Comparative Analysis with Baseline Models

To rigorously evaluate the performance of the proposed method, a comparative study is conducted against three established baseline models: Back Propagation Neural Network (BPNN), Support Vector Regression (SVR), and the D-LSTM. The former two represent classic static machine learning methods, while the latter serves as a deep learning benchmark without the spatiotemporal partitioning strategy.

The quantitative comparison results on the test set are summarized in Table 7. Observation reveals that the proposed TP-D-LSTM outperforms all baseline models across all evaluation metrics.

Table 7. Performance comparison.

Models	RMSE	MAE	R^2
BPNN	25.76	15.56	0.8356
SVR	25.92	13.34	0.8336
D-LSTM	13.83	8.43	0.9550
TP-D-LSTM	6.15	4.43	0.9812

As shown in Table 7, traditional static models (BPNN and SVR) fail to capture temporal dependencies effectively, resulting in RMSE values exceeding 25. While the D-LSTM utilizes memory mechanisms to reduce the RMSE to 13.83, it still struggles with the high heterogeneity of traffic conditions, yielding an MAE of 8.43. In contrast, the proposed method achieves a substantial further reduction in error. Specifically, the RMSE drops to 6.15 and the MAE to 4.43, representing an improvement of over 50% compared to the global baseline. The R^2 value of 0.981 further confirms its superior goodness of fit.

In addition to statistical metrics, we analyzed the stability of the models through the error probability distribution, as shown in Figure 13. The histogram displays the probability density of prediction errors for each model.

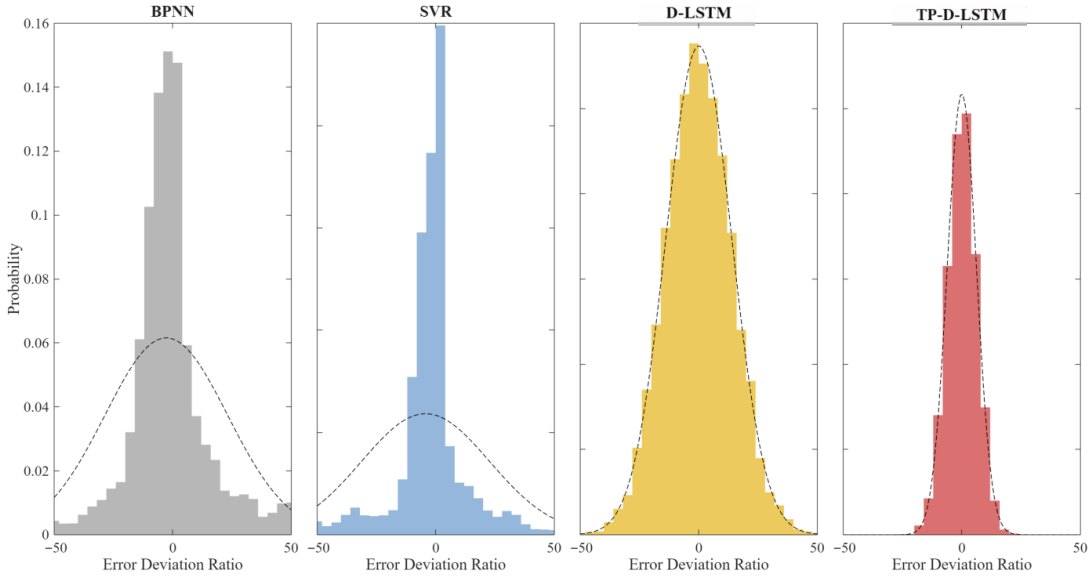


Figure 13. Error distribution of the prediction results.

The analysis highlights the robustness of TP-D-LSTM through error probability distributions. The model exhibits a sharp, high-kurtosis peak around zero, indicating consistently low deviations across all samples. This stability is particularly relevant for practical electric bus operations, supporting reliable energy estimation and informed decisions for charging scheduling and fleet management. This high kurtosis characteristic verifies the model's strong robustness and stability, demonstrating its ability to maintain precise control even under dynamic and volatile traffic conditions.

6. Conclusions

To address the limitations of insufficient dynamic feature representation and the lack of physical interpretability in traditional electric bus energy prediction models, this paper proposes a Time-Partitioned Dual-Layer LSTM framework. By implementing a rule-based soft-partitioning strategy, the proposed method effectively decouples complex urban driving cycles into distinct operational regimes, thereby mitigating the "pattern averaging" problem inherent in global models. Simultaneously, the physics-constrained dual-layer architecture bridges the gap between data-driven fitting and physical fidelity by explicitly integrating electrochemical state prediction with theoretical formula constraints. Experimental validation on real-world cloud-logged data from the Jinan K260 bus line demonstrates that the proposed framework significantly outperforms traditional baselines, achieving a predicted RMSE of 6.15 even in the most volatile evening peak scenarios. These findings confirm that the integration of explicit scenario isolation and physical priors

provides a robust, interpretable, and high-precision solution for the energy management and range estimation of electric vehicles in complex urban environments.

For future work, we plan to improve robustness to sensor failures or missing data, extend the TP-D-LSTM to multiple bus lines, and integrate SOC prediction into fleet management. Furthermore, we intend to conduct a more systematic evaluation of the time-partitioning strategy, specifically focusing on its impact on model stability and potential overfitting risks under a wider range of traffic and environmental conditions. Additionally, future research could leverage advanced simulation assets tailored to transit scenarios, such as the KIT Bus shuttle model for the CARLA simulator [31], to provide a controlled environment for validating the TP-D-LSTM framework under diverse edge-case traffic conditions.

In addition, expanding the dataset and conducting broader comparative evaluations represent an important direction for future research. By including more diverse operational routes, vehicle types, and prediction targets, these efforts will enable a fair and systematic assessment of deep learning architectures for electric bus energy prediction.

Author Contributions: Conceptualization, Y.W. (Yue Wang), W.Z. and B.W.; methodology, Y.W. (Yue Wang), Y.W. (Yu Wang) and W.Z.; software, Y.W. (Yu Wang); validation, S.L., Y.Z., G.Y. and Y.L.; formal analysis, Y.W. (Yu Wang) and Y.W. (Yue Wang); investigation, Y.W. (Yu Wang) and S.L.; resources, B.W., S.L., Y.Z. and G.Y.; data curation, Y.W. (Yu Wang), S.L., Y.Z. and Y.L.; writing—original draft preparation, Y.W. (Yu Wang); writing—review and editing, Y.W. (Yue Wang), W.Z. and B.W.; visualization, Y.W. (Yu Wang); supervision, Y.W. (Yue Wang), W.Z. and B.W.; project administration, B.W. and W.Z.; funding acquisition, B.W. and Y.W. (Yue Wang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Projects for Urgently Needed Talents in Key Support Regions of Shandong Province and the Fundamental Research Funds for the Central Universities (FRF-TP-24-019A, QNXM20250006).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Yue Wang, Shiqi Liu, Yanpeng Zhu, Bo Wang, and Guoqun Yao were employed by the company Zhongtong Bus Holding Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Chen, Y.; Zhang, Y.; Sun, R. Data-driven estimation of energy consumption for electric bus under real-world driving conditions. *Transp. Res. Part D Transp. Environ.* **2021**, *98*, 102969. [CrossRef]
2. Zhang, J.; Wang, Z.; Liu, P.; Zhang, Z. Energy consumption analysis and prediction of electric vehicles based on real-world driving data. *Appl. Energy* **2020**, *275*, 115408. [CrossRef]
3. Logan, K.G.; Nelson, J.D.; Hastings, A. Electric and hydrogen buses: Shifting from conventionally fuelled cars in the UK. *Transp. Res. Part D Transp. Environ.* **2020**, *85*, 102350. [CrossRef]
4. Li, P.; Zhang, Y.; Zhang, Y.; Zhang, Y.; Zhang, K. Prediction of electric bus energy consumption with stochastic speed profile generation modelling and data driven method based on real-world big data. *Appl. Energy* **2021**, *298*, 117204. [CrossRef]
5. Çiçek, A.; Karakuş, F.; Erdinç, O. A novel concept for optimal operation of multi-swap stations serving electric metrobuses. *Energy Sustain. Dev.* **2025**, *86*, 101722. [CrossRef]
6. Lim, L.K.; Muis, Z.A.; Ho, W.S.; Hashim, H.; Bong, C.P.C. Review of the energy forecasting and scheduling model for electric buses. *Energy* **2023**, *263*, 125773. [CrossRef]
7. Aljohani, T.M.; Ebrahim, A.; Mohammed, O. Real-time metadata-driven routing optimization for electric vehicle energy consumption minimization using deep reinforcement learning and Markov chain model. *Electr. Power Syst. Res.* **2021**, *192*, 106962. [CrossRef]
8. Zhang, Z.; Ye, B.; Wang, S.; Ma, Y. Analysis and estimation of energy consumption of electric buses using real-world data. *Transp. Res. Part D Transp. Environ.* **2024**, *126*, 104017. [CrossRef]

9. Szilassy, P.Á.; Földes, D. Consumption estimation method for battery-electric buses using general line characteristics and temperature. *Energy* **2022**, *261*, 125080. [CrossRef]
10. Nan, S.; Tu, R.; Li, T.; Sun, J.; Chen, H. From driving behavior to energy consumption: A novel method to predict the energy consumption of electric bus. *Energy* **2022**, *261*, 125188. [CrossRef]
11. Ma, X.; Miao, R.; Wu, X.; Liu, X. Examining influential factors on the energy consumption of electric and diesel buses: A data-driven analysis of large-scale public transit network in Beijing. *Energy* **2021**, *216*, 119196. [CrossRef]
12. Ma, X.; Tao, Z.; Wang, Y.; Yu, H.; Wang, Y. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp. Res. Part C Emerg. Technol.* **2015**, *54*, 187–197. [CrossRef]
13. Jiang, J.; Yu, Y.; Min, H.; Cao, Q.; Sun, W.; Zhang, Z.; Luo, C. Trip-level energy consumption prediction model for electric bus combining Markov-based speed profile generation and Gaussian processing regression. *Energy* **2023**, *263*, 125866. [CrossRef]
14. Abdelaty, H.; Mohamed, M. A prediction model for battery electric bus energy consumption in transit. *Energies* **2021**, *14*, 2824. [CrossRef]
15. Fiori, C.; Ahn, K.; Rakha, H.A. Power-based electric vehicle energy consumption model: Model development and validation. *Appl. Energy* **2016**, *168*, 257–268. [CrossRef]
16. Pamuła, T.; Pamuła, W. Estimation of the energy consumption of battery electric buses for public transport networks using real-world data and deep learning. *Energies* **2020**, *13*, 2340. [CrossRef]
17. Dong, C.; Xiong, Z.; Zhang, C.; Li, N.; Li, Y.; Xie, N.; Zhang, J.; Wang, H. A transformer-based approach for deep feature extraction and energy consumption prediction of electric buses based on driving distances. *Appl. Energy* **2025**, *380*, 123941. [CrossRef]
18. Li, Z.; Cui, Z.; Liao, H.; Ash, J.; Zhang, G.; Xu, C.; Wang, Y. Steering the future: Redefining intelligent transportation systems with foundation models. *CHAIN* **2024**, *1*, 46–53. [CrossRef]
19. Wang, H.; Song, J.; Li, X. A survey of electric vehicles in sustainable systems: Environmental benefits, pattern analysis, and AI-driven operation methodologies. *CHAIN* **2024**, *1*, 177–202. [CrossRef]
20. Zhao, J.; He, J. Energy consumption prediction for electric buses based on traction modeling and LightGBM. *World Electr. Veh. J.* **2025**, *16*, 159. [CrossRef]
21. Ng, M.-F.; Zhao, J.; Yan, Q.; Conduit, G.J.; Seh, Z.W. Predicting the state of charge and health of batteries using data-driven machine learning. *Nat. Mach. Intell.* **2020**, *2*, 161–170. [CrossRef]
22. Zhang, W.; Zhang, H.; Bi, Z. A time series physics-informed neural network framework for the state of health estimation of lithium-ion batteries. *J. Energy Storage* **2026**, *142*, 119549. [CrossRef]
23. Yu, H.; Zhang, H.; Zhang, Z.; Yang, S. State estimation of lithium-ion batteries via physics-machine learning combined methods: A methodological review and future perspectives. *eTransportation* **2025**, *24*, 100420. [CrossRef]
24. Xu, H.; Tu, R.; Li, T.; Chen, H. Interpretable bus energy consumption model with minimal input variables considering powertrain types. *Transp. Res. Part D Transp. Environ.* **2023**, *119*, 103742. [CrossRef]
25. Mersha, M.; Lam, K.; Wood, J.; AlShami, A.K.; Kalita, J. Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction. *Neurocomputing* **2024**, *599*, 128111. [CrossRef]
26. Mei, P.; Karimi, H.R.; Chen, F.; Yang, S.; Huang, C.; Qiu, S. A learning-based vehicle-cloud collaboration approach for joint estimation of state-of-energy and state-of-health. *Sensors* **2022**, *22*, 9474. [CrossRef] [PubMed]
27. Wang, Y.; Tang, H.; Pan, B.; Wang, S.; Niu, Y.; Wang, R. A cloud energy management strategy for intelligent connected HEVs based on end-edge-cloud collaboration. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2024**, *239*, 4756–4773. [CrossRef]
28. Guo, S.; Zhao, C. iEVEM: Big data-empowered framework for intelligent electric vehicle energy management. *Systems* **2025**, *13*, 118. [CrossRef]
29. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
30. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [CrossRef]
31. Xiang, Y.; Wang, S.; Su, T.; Li, J.; Mao, S.S.; Geimer, M. KIT Bus: A Shuttle Model for CARLA Simulator. In *Proceedings of the 2021 IEEE Industrial Electronics and Applications Conference (IEACon), Penang, Malaysia, 22–23 November 2021*; IEEE: New York, NY, USA, 2021; pp. 7–12. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.