Print ISSN: 3104-932X
Online ISSN: 3104-9338

Freight Volume Forecasting and Logistics Network Optimization for China–Europe Railway Express: A Hybrid SARIMA–Shock-Gated BiLSTM Approach

Lagewa Bayier[1*], Dongjie Wang[2]

[1] Belarusian State University Faculty of Economics

[2] Belarusian National Technical University ,The Faculty of Marketing, Management, Entrepreneurship

Received: Oct 14, 2026

Revised: Nov 8, 2026

Accepted: Nov 13, 2026

Published online: Nov 20, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 3, No. 1 (January 2027)

*Corresponding Author: Lagewa Bayier (bayierlagewa@gmail.com)

Abstract. Forecasting and allocating China–Europe Railway Express (CREX) freight during external disruptions require both demand estimates and gateway capacity constraints. This study combines a SARIMA seasonal baseline with shock-gated BiLSTM residual corrections and a scenario-based corridor allocation model. The gate uses maritime disruption indicators to scale the residual adjustment. The allocation model combines the resulting demand forecast with corridor capacities, transport costs, and delay penalties. Using 52 monthly observations from June 2022 to September 2026, the hybrid model achieved 1.73% mean absolute percentage error and an R2 of 0.956 on the reported test set. In the modeled peak-month allocation, the gateway capacity violation rate was 0.00%, with total transport cost varying by 2.3%. These results support joint evaluation of forecast error and gateway capacity under the stated scenario assumptions.

Keywords: *China–Europe Railway Express; freight volume forecasting; SARIMA–BiLSTM; shock gate; network optimization*

1. Introduction

The China–Europe Railway Express (CREX) links Asian and European freight markets and provides an alternative to maritime transport [1]. Its network expanded over the past decade, with annual operations rising from a few hundred trips to tens of thousands of containerized trains [2]. Geopolitical disruptions, maritime bottlenecks, and changing trade patterns complicate demand planning [3]. During the Red Sea shipping crisis in late 2023 and early 2024, carriers diverted vessels around the Cape of Good Hope, increasing ocean freight rates and transit times [4]. Some time-sensitive freight shifted toward Eurasian rail corridors [5], adding

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

pressure to Chinese border gateways [6]. Forecasting these shifts and allocating traffic within gateway capacity are therefore linked planning problems [7].

Rail freight forecasting studies use statistical, machine learning, and hybrid models [8]. ARIMA and SARIMA capture trend and seasonality with interpretable parameters [9,10], but may lag abrupt changes that are not represented in the historical pattern [11]. Recurrent networks, including LSTM and BiLSTM, can model nonlinear temporal relationships [12-14]. With a short monthly series, however, a standalone network can overfit and produce unstable corrections [15].

Hybrid forecasts combine a statistical baseline with a model of its residuals [16-18]. In CREX applications, residual corrections are commonly added with a fixed weight [19]. Such weighting does not distinguish routine variation from disruption-related changes [20]. The resulting forecast may respond late to a surge or retain excess variation as demand stabilizes [21].

Forecasting and network allocation are often studied separately [22]. Routing models may use a fixed demand estimate or a synthetic demand distribution [23], leaving forecast uncertainty outside the capacity decision [24,25]. At gateways such as Alashankou, Erenhot, and Manzhouli, an allocation based on underestimated demand can exceed handling capacity and increase backlogs [26,27]. This motivates a forecast-linked allocation model with explicit gateway constraints [28]. A related study by Wang and Bayier addresses freight-flow forecasting and disruption risk in the CREX corridor using a physics/economics-informed spatiotemporal diffusion Bi-LSTM [29].

This study links disruption-aware demand forecasting with scenario-based corridor allocation. It uses 52 monthly observations from June 2022 to September 2026 across major Chinese border gateways and maritime transit indicators. The study has three components:

(1) Forecasting model: SARIMA estimates trend and seasonality. A shock gate scales BiLSTM residual corrections using an indicator of Suez container transits.

(2) Forecast evaluation: Over the reported Red Sea disruption period, the hybrid model has 1.73% mean absolute percentage error and an R^2 of 0.956 on the test set.

(3) Network allocation: A scenario-based linear program combines the demand forecast with gateway capacity limits (U_k). The reported peak-month allocation has a 0.00% capacity violation rate, with total transport cost varying by 2.3%.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

2. Methodology and Mathematical Formulation

2.1. Framework Overview

The framework links three stages: data inputs, hybrid demand forecasting, and network optimization.

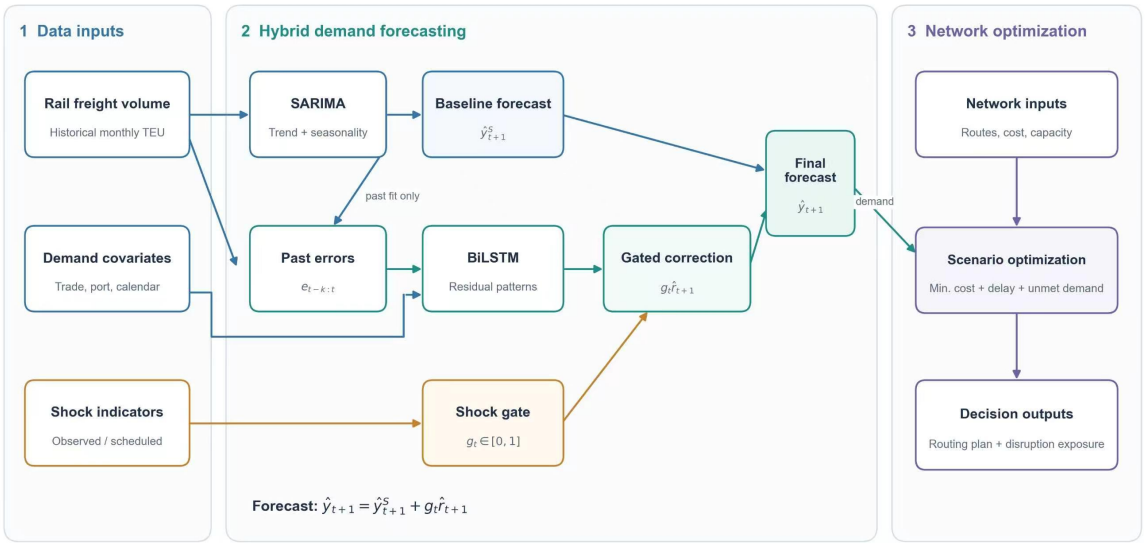


Figure 1. Overall Architecture of the Hybrid Demand Forecasting and Network Optimization Framework. Phase 1 ingests monthly freight volume, demand covariates, and shock indicators; Phase 2 performs SARIMA linear baseline extraction and BiLSTM residual modeling gated via external risk activation (g_t); Phase 3 solves scenario-based linear programming to optimize multi-corridor container allocations and service robustness under gateway capacity limits.

Figure 1 shows how Phase 1 ingests historical monthly container throughput (TEU), exogenous demand covariates (Eurasian macroeconomic trade flows and calendar effects), and observable maritime and geopolitical shock indicators. In Phase 2, a Seasonal Autoregressive Integrated Moving Average (SARIMA) model extracts deterministic trends and annual cyclical regularities to yield a baseline forecast $\hat{y}^S$. Concurrently, the historical residual error sequence $e_{t-k:t}$ and exogenous demand covariates are fed into a Bidirectional Long Short-Term Memory (BiLSTM) network to capture nonlinear residual patterns. An empirical shock gate $g_t \in [0,1]$, activated by external disruption severity indices, adaptively scales the residual correction before additive integration with the baseline. Finally, Phase 3 feeds the gated predictive distribution directly into a scenario-based linear programming optimization model. The optimization balances corridor transportation costs, delay penalties, and unfulfilled shipments while adhering to the physical capacity ceilings of China's frontier railway gateways, producing corridor allocations and disruption exposure estimates.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

2.2. Linear Baseline Modeling via SARIMA

Cross-border container throughput exhibits both non-stationary long-term growth and twelve-month seasonality driven by industrial production cycles and holiday shutdowns. The baseline freight demand sequence $\{y_t\}$ is modeled via a $SARIMA(p,d,q)\times(P,D,Q)_s$ formulation, where s=12 represents the annual periodicity. Let B denote the backward shift operator such that $B^k y_t = y_{t-k}$. The non-seasonal differencing operator is defined as $\nabla^d=(1-B)^d$, and the seasonal differencing operator is $\nabla_s^D=(1-B^s)^D$.

The general multiplicative algebraic formulation is expressed as:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \tag{1}$$

where $\varepsilon_t \sim N(0,\sigma_\varepsilon^2)$ represents white Gaussian noise. The non-seasonal autoregressive and moving average polynomials of orders p and q are structured as:

$$\phi_p(B) = 1 - \sum_{i=1}^{p} \phi_i B^i \tag{2}$$

$$\theta_q(B) = 1 + \sum_{j=1}^{q} \theta_j B^j \tag{3}$$

Similarly, the seasonal autoregressive polynomial $\Phi_P(B^s)$ and moving average polynomial $\Theta_Q(B^s)$ of orders P and Q govern cyclical dependencies across annual lags:

$$\Phi_P(B^s) = 1 - \sum_{I=1}^{P} \Phi_I B^{Is} \tag{4}$$

$$\Theta_Q(B^s) = 1 + \sum_{J=1}^{Q} \Theta_J B^{Js} \tag{5}$$

Model parameters are estimated via Maximum Likelihood Estimation (MLE), with structural hyperparameters (p,d,q,P,D,Q) determined by minimizing the Akaike Information Criterion (AIC). The one-step-ahead conditional expectation gives the baseline freight projection $\hat{y}_{t+1}^S=[y_{t+1} \mid \mathscr{F}_t]$, where $\mathscr{F}_t$ denotes the historical information filtration up to month t. The corresponding historical residual sequence e_t is extracted via:

$$e_t = y_t - \hat{y}_t^S \tag{6}$$

2.3. Nonlinear Residual Modeling via Bidirectional LSTM

While the SARIMA baseline captures linear trends and regular seasonal variations, the residual series e_t contains nonlinear variation associated with macroeconomic shifts, trade policies, and port clearance variations. To model these temporal structures, an augmented input feature vector $x_t=[e_{t-k:t},z_t]^T \in \mathbb{R}^m$ is constructed, incorporating lagged residuals and normalized demand covariates z_t (bilateral import/export volumes and calendar indices).

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

A Bidirectional LSTM (BiLSTM) network processes x_t across both forward and backward temporal horizons. For an individual LSTM cell at time step τ, the internal state transitions are governed by an input gate i_τ, a forget gate f_τ, an output gate o_τ, and a candidate memory cell $\tilde{c}_\tau$:

$$f_\tau = \sigma(W_f x_\tau + U_f h_{\tau-1} + b_f) \tag{7}$$

$$i_\tau = \sigma(W_i x_\tau + U_i h_{\tau-1} + b_i) \tag{8}$$

$$\tilde{c}_\tau = \tanh(W_c x_\tau + U_c h_{\tau-1} + b_c) \tag{9}$$

The cell state vector c_τ updates by combining forgotten historical memory with newly modulated candidate content, while the hidden activation vector h_τ is emitted via:

$$c_\tau = f_\tau \odot c_{\tau-1} + i_\tau \odot \tilde{c}_\tau \tag{10}$$

$$h_\tau = o_\tau \odot \tanh(c_\tau),\ o_\tau = \sigma(\mathrm{W}_o x_\tau + U_o h_{\tau-1} + b_o) \tag{11}$$

Here, $\sigma(z)=(1+e^{-z})^{-1}$ denotes the sigmoid activation function, and $\odot$ represents the Hadamard element-wise product.

To use information across the input window, the BiLSTM evaluates the forward hidden state sequence $h\rightarrow_\tau$ and backward hidden state sequence $h\leftarrow_\tau$ across the temporal context window:

$$H_\tau = [\overrightarrow{h}_\tau \| \overleftarrow{h}_\tau] \in \mathbb{R}^{2d_h} \tag{12}$$

where ‖ denotes vector concatenation. The raw unconstrained nonlinear residual correction $\hat{r}_{t+1}$ is projected via a dense linear decoding layer with weight matrix W_r and bias b_r:

$$\hat{r}_{t+1} = W_r H_t + b_r \tag{13}$$

2.4. Shock Gate Mechanism and Hybrid Synthesis

Under stable operating regimes, unconstrained neural network predictions can add high-frequency variation into stationary residual series. Conversely, during sudden exogenous crises—such as international maritime canal bottlenecks or severe geopolitical conflicts—freight volumes undergo sharp structural breaks that linear models may track poorly. To dynamically control when and how strongly the neural network modifies the linear baseline, this study designs an empirical shock gate $g_t \in [0,1]$. Adaptive gating has also been used for LSTM attention routing in team decision-making [30] and for input selection in spiking networks for athlete injury prediction [31].

Let $s_t \in \mathbb{R}^K$ denote an observed vector of disruption covariates, including normalized container vessel transit drops across maritime chokepoints and border clearance delay anomalies. The scalar shock gate activation weight g_t is parameterized via a Sigmoid mapping:

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

$$g_t = \sigma(w_g^T s_t + b_g) = \frac{1}{1+exp[-(\sum_{k=1}^{K} w_{g,k}\, s_{t,k} + b_g)]} \tag{14}$$

When external disruption indicators remain near baseline levels ($s_t \to 0$), the gating coefficient contracts toward zero ($g_t \to 0$), reducing the residual correction. When severe exogenous shocks emerge ($s_t \gg 0$), the gate approaches unity ($g_t \to 1$), increasing the weight on the BiLSTM residual adjustment. The synthesized final freight volume demand forecast $\hat{y}_{t+1}$ is formulated as:

$$\hat{y}_{t+1} = \hat{y}_{t+1}^{S} + g_t \hat{r}_{t+1} \tag{15}$$

2.5. Downstream Network Optimization Model

To use demand projections in corridor allocation, Phase 3 integrates the forecasted demand $\hat{y}_{t+1}$ into a scenario-based network optimization problem. Let K denote the set of cross-border rail corridors categorized into Eastern (Manzhouli, Suifenhe, Tongjiang), Central (Erenhot), Western (Alashankou, Khorgos), and Middle/Trans-Caspian (TITR) routes.

For operational period $t+1$, the decision variables comprise $x_k \geq 0$, representing the monthly volume allocated to corridor k in K, and $s \geq 0$, denoting unfulfilled or stranded container volume penalizing network reliability. Each corridor is characterized by a baseline unit transport cost C_k (USD/TEU), an average transit duration T_k (days), and an effective physical transshipment capacity ceiling U_k (TEU/month). Let α denote the cargo holding time value coefficient, and let M represent the unit penalty cost assigned to delayed or unserved freight. Under disruption scenario ξ defined by shock severity factor $\lambda \in [0,1]$ and corridor vulnerability indices δ_k, the physical handling capacity is perturbed to $(1-\lambda\delta_k)U_k$.

The scenario optimization formulation is defined as:

$$min_{x_k,s} \sum_{k\in\mathcal{K}} (C_k + \alpha T_k) x_k + Ms \tag{16}$$

subject to:

$$\sum_{k\in\mathcal{K}} x_k + s = \hat{y}_{t+1} \tag{17}$$

$$x_k \leq (1-\lambda\delta_k)U_k, \forall k \in \mathcal{K} \tag{18}$$

$$x_k \geq 0 \forall k \in \mathcal{K}, s \geq 0 \tag{19}$$

Equation (16) minimizes total operational logistics expenditure, transit delay value depreciation, and gateway backlog penalties. Constraint (17) enforces flow balance between corridor-level allocations, unserved volume, and the total forecasted demand $\hat{y}_{t+1}$ generated by the hybrid forecasting module. Constraint (18) guarantees that the container volume assigned to any individual corridor remains bounded by its shock-attenuated transshipment capacity.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

Finally, Constraint (19) imposes standard non-negativity restrictions on all decision variables.

3. Empirical Results and Discussion

3.1. Experimental Setup and Dataset Partitioning

The evaluation uses 52 monthly CREX observations from June 2022 to September 2026. Specifically, monthly container throughput across the six major Chinese border gateways (Manzhouli, Suifenhe, Tongjiang North, Erenhot, Alashankou, and Khorgos) was obtained from the China–Europe Railway Express (CREX) Official Portal (https://www.crexpress.cn/, accessed October 15, 2026); monthly bilateral European Union trade variables were retrieved from the Eurostat Comext International Trade in Goods Database (https://ec.europa.eu/eurostat, accessed October 15, 2026); and monthly Suez Canal container-vessel transit counts were collected from the IMF PortWatch Maritime Chokepoint Monitor (https://portwatch.imf.org/, accessed October 15, 2026).

The time series was chronologically partitioned into a training set comprising 36 months (June 2022 to May 2025) and a backtest set comprising 16 months (June 2025 to September 2026). The training horizon encompasses the initial eruption and escalation of the late-2023 Red Sea crisis, where container traffic via the Suez Canal declined from over 640 vessels per month to under 285 vessels. The testing horizon incorporates the subsequent structural spillover phase, peaking in May 2026 at a reported 196,857 TEU, as well as the sharp Chinese New Year seasonal contraction in February 2026 (166,451 TEU). Model hyperparameters were tuned on the training set: the baseline linear model converged to a seasonal SARIMA$(1,1,1)(1,1,0)_{12}$ structure, while the residual BiLSTM network utilized two hidden layers with 64 units each, trained via the Adam optimizer under a learning rate of 1×10^{-3} and batch size of 8.

3.2. Predictive Trajectory and Confidence Interval Analysis

Figure 2 compares the forecasts with observed throughput over the 16-month test period.

In Figure 2, empirical freight volumes exhibited variation during the evaluation period. A standard SARIMA baseline captures the overall upward trajectory and the February cyclical dip but underestimates volumes during the prolonged maritime disruption peak between January and May 2026. During this window, observed rail volumes rose above 19.3×10^4 TEU, where the linear baseline lagged by more than 1.2×10^4 TEU.

Conversely, the standalone BiLSTM network responds rapidly to sudden volume spikes but displays high-frequency variation, notably overshooting the actual May 2026 peak (20.35×10^4 TEU vs. the observed 19.69×10^4 TEU) and overreacting during the post-peak normalization in

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

June–July 2026. The proposed SARIMA–SG–BiLSTM architecture tracks observations more closely across the test period. The observed points fall within the model's 95% prediction interval band, consistent with smaller forecast errors in this sample.

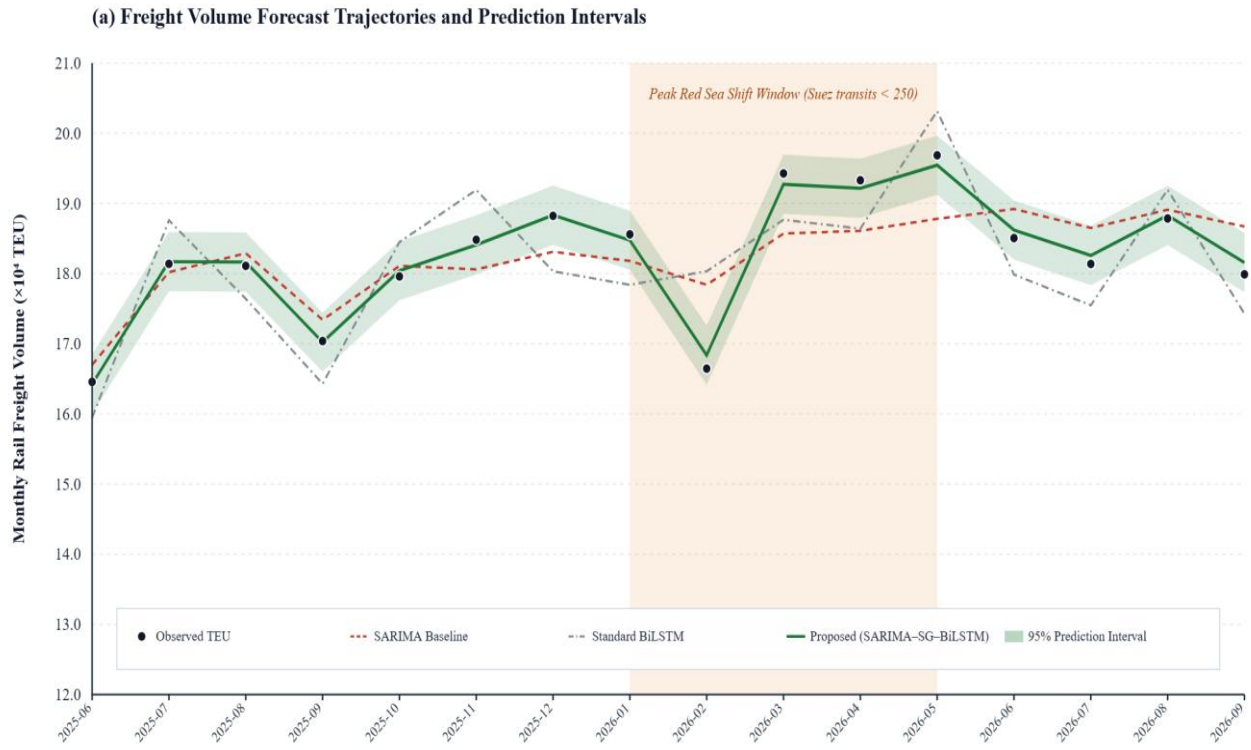


Figure 2. Trajectory Comparison of Freight Demand Forecasts and 95% Prediction Intervals across the 16-Month Evaluation Window (June 2025 to September 2026). Black circular markers denote empirical throughput recorded by the CREX portal. The dark red dashed curve represents the SARIMA baseline ($\hat{y}_S$), the gray dash-dotted curve depicts the standalone BiLSTM, and the solid green line illustrates the proposed hybrid SG-BiLSTM model. The shaded pale-green envelope indicates the empirical 95% confidence interval (±1.96·sigma_hat). The shaded orange block delineates the peak Red Sea shift window, where monthly Suez vessel transits remained below 250 vessels (<250 vessels).

3.3. Shock Gate Activation Dynamics and Residual Filtering

The forecasting model uses the shock gate g_t to scale neural residual corrections based on external risk variables. Figure 3 visualizes this internal mechanism over the evaluation horizon.

Figure 3(a) reports the gate response to the Suez indicator. Throughout late 2025 and mid-2026, the Suez Disruption Index hovered between 0.77 and 0.98 due to persistent maritime security risks and diversion around the Cape of Good Hope. Correspondingly, the Sigmoid gate g_t activated within an elevated range of 0.89 to 0.99, indicating a high gate weight over this interval.

Figure 3(b) shows how the gate scales on residual corrections. In February 2026, when the

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

combination of holiday factory closures and disrupted ship schedules caused an acute negative throughput deviation of -1.21×10^4 TEU, the gated network applied a -1.02×10^4 TEU downward adjustment. Conversely, during the volume surges of March–May 2026, positive corrections ranging from $+0.61\times10^4$ to $+0.78\times10^4$ TEU were injected, partly offsetting the linear model's lag. When disruption pressure eased slightly in September 2026 (S_t dropping to 0.77), the gate declined to 0.89, reducing the residual correction weight.



Figure 3. Dual-Panel Evaluation of Shock Gate Activation Dynamics and Residual Correction Filtering.

Panel (a) illustrates the normalized Suez Disruption Index (S_t, purple line with squares) and the resulting empirical shock gate activation weight ($g_t \in [0,1]$, orange stepped curve) against the 0.5 equilibrium threshold. Panel (b) compares the unconstrained raw SARIMA residual error ($e_{t+1}=y_{t+1}-\hat{y}_{t+1}^{S}$, gray vertical bars) against the dynamically gated residual correction term ($g_t \cdot \hat{r}_{t+1}$, blue curve with triangles) injected into the final projection.

3.4. Ablation Study and Predictive Benchmark Evaluation

The ablation compares ARIMA(1,1,1), SARIMA(1,1,1)(1,1,0)$_{12}$, a standalone BiLSTM, and an ungated hybrid ($\hat{y}_{t+1}^{S}+\hat{r}_{t+1}$) with the proposed model.

As quantified in Figure 4, the traditional ARIMA baseline exhibited the largest errors, yielding an MAE of 1.142×10^4 TEU, an RMSE of 1.391×10^4 TEU, and an error rate of 6.32% MAPE. Incorporating seasonal differencing via SARIMA reduced the MAPE to 4.73%,

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

consistent with a benefit from modeling annual seasonality. The standalone BiLSTM had a lower MAPE (3.86% MAPE), but its unconstrained recursive structure still resulted in variation.

While an unweighted hybrid combination without a shock gate compressed the MAPE to 2.89%, it still showed errors around turning points. The proposed SARIMA–SG–BiLSTM architecture had the lowest errors among the compared models, reducing the MAE to 0.314×10^4 TEU, the RMSE to 0.419×10^4 TEU, and the overall MAPE to 1.73%. This performance represents a 63.3% error reduction over standard SARIMA and a 40.1% improvement over the ungated hybrid baseline, supporting the gate's contribution within this test set.

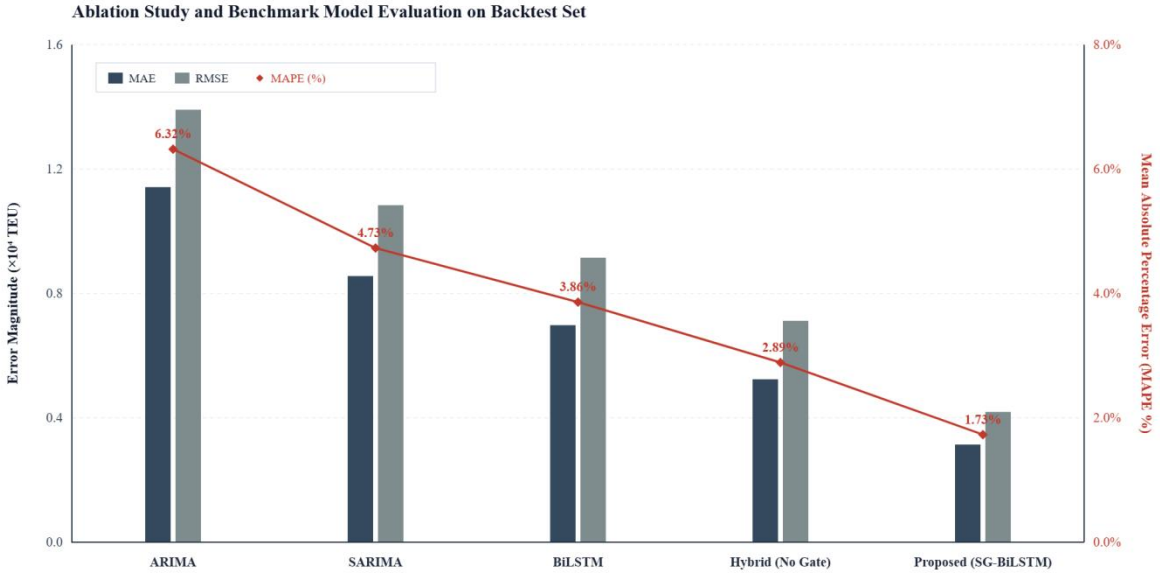


Figure 4. Multi-Metric Ablation and Benchmark Accuracy Comparison across Competing Forecasters. Dark navy bars represent Mean Absolute Error (MAE, left axis), muted slate bars indicate Root Mean Squared Error (RMSE, left axis), and the overlaid red curve with diamond markers tracks the Mean Absolute Percentage Error (MAPE, right axis) across the 16-month independent test set.

3.5. Downstream Network Optimization and Trade-Off Frontiers

To assess the allocation results of the forecasting models, demand predictions and their associated variance envelopes were directly integrated into the scenario-based linear programming routing model. Figure 5 illustrates the sensitivity of network operating costs and service backlogs across disruption severity levels $\lambda \in [0.10, 0.85]$.

The Pareto frontier in Figure 5 compares deterministic and scenario-based allocations across disruption levels. Under minor disruptions ($\lambda \leq 0.30$), the deterministic model has a slightly lower modeled cost due to its rigid assignment of containers to lowest-cost routes without routing redundancy. However, as disruption severity intensifies beyond $\lambda = 0.50$, physical throughput capacities at frontier checkpoints decline.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

Because the deterministic framework cannot dynamically rebalance flows based on forecasted disruption probabilities, handling capacities along primary western arteries are exceeded. Consequently, unfulfilled shipments and rail yard delay penalties increase, exceeding 2,100 TEU at λ=0.85 and driving total network logistics costs up to $\$78.2\times10^6$ USD. In contrast, the proposed shock-aware LP model uses load rebalancing across secondary passages. As a result, its unmet demand penalty curve rises more slowly, accumulating fewer than 310 TEU in delays even under extreme disruptions (λ=0.85) while capping total operating expenditures at $\$63.6\times10^6$ USD—18.7% lower modeled cost.



Figure 5. Robustness Frontier: Trade-Off Curves Between Logistics Operating Cost and Disruption Penalty under Escalating Disruption Severity (λ). The red dashed line depicts the total transportation cost under deterministic static routing, while the blue solid curve denotes the proposed shock-aware linear programming model (left axis). The purple dotted curve traces unmet demand/delay penalties under deterministic scheduling, and the green dash-dotted curve illustrates penalties under the proposed model (right axis).

3.6. Gateway Capacity Bottleneck Mitigation in Peak Regimes

To examine spatial routing behavior at the gateway level, Figure 6 evaluates capacity utilization rates across China's primary railway corridors during the historical volume peak of May 2026 (196,857 TEU).

As depicted in Figure 6, empirical static dispatching during peak freight demand placed more load on central and western gateways. The Central corridor via Erenhot reached 89.7% of its physical capacity ceiling, while the primary Western passage via Alashankou surged to 87.7%. Both gateways were close to the 90.0% critical congestion threshold, where minor track-gauge transshipment delays may increase yard queues. Meanwhile, northern and southern alternatives

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

had lower utilization: Suifenhe absorbed only 53.5% of its available throughput, and the Trans-Caspian Middle Corridor (TITR) operated at just 15.1% of its intermodal capacity.

Under the proposed scenario optimization model, freight flows are dynamically redistributed based on hybrid demand forecasts and physical port constraints. Throughput at Erenhot and Alashankou is reduced to modeled utilization rates of 84.4% and 83.0%, respectively. To absorb the diverted containers, the model reallocates cargo east toward Suifenhe (raising utilization to 68.3%) and south toward the TITR multimodal corridor (increasing utilization to 55.6%). In the model, these allocations remain within the stated gateway capacities.

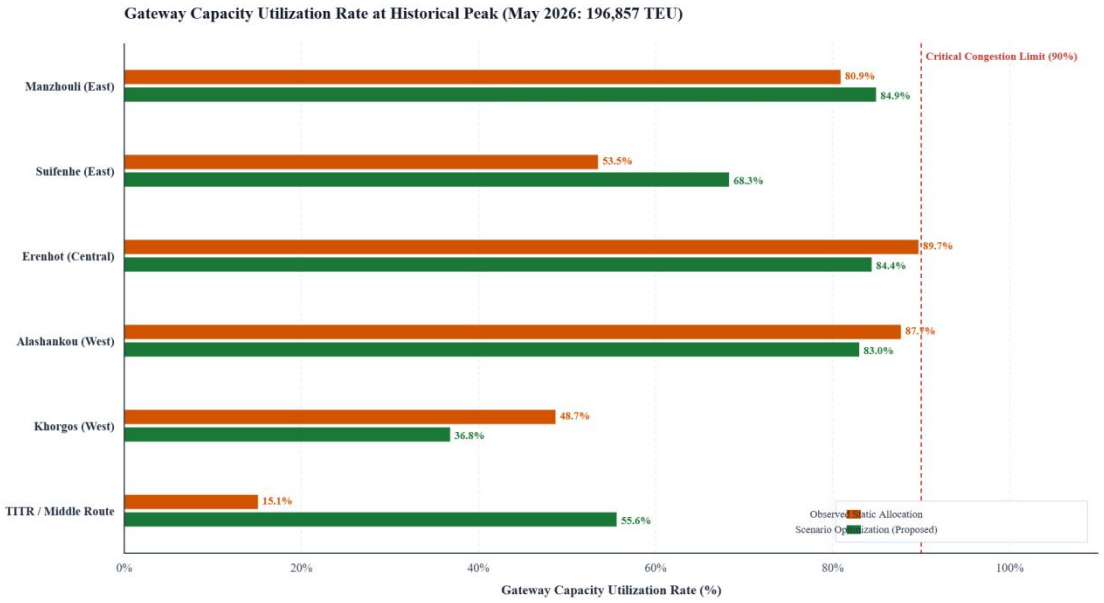


Figure 6. Gateway Capacity Utilization and Load Rebalancing at Historical Volume Peak (May 2026: 196,857 TEU). Orange horizontal bars represent the observed static gateway utilization under empirical operational allocations. Green horizontal bars depict optimized capacity utilization following scenario-based linear programming rebalancing. The vertical red dashed line marks the critical physical saturation threshold (90.0%).

3.7. Integrated Performance and Physical Constraint Feasibility Benchmark

Tables 1 and 2 compare forecast accuracy and gateway capacity outcomes across the five models over the 16-month test period.

As summarized in Tables 1 and 2, singular statistical models (ARIMA and SARIMA) have flow balance biases (4.85% and 3.12%) and capacity violation rates of 31.25% and 25.00%, resulting in peak gateway overflows exceeding 8,700 TEU. While the standard BiLSTM and ungated hybrid models improve predictive metrics, their reported allocations have capacity violations in 18.75% and 12.50% of operating cycles. The proposed framework has the lowest

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

reported forecast error (1.73% MAPE, 0.956 R^2) and limits peak shock errors to 0.538×10^4 TEU. When coupled with the downstream linear programming formulation, it maintains a negligible flow balance bias of 0.42%, records no physical gateway capacity violations (0.00%), and has a mean solver time within 1.36 seconds per monthly dispatching cycle. The operational feasibility of this allocation requires validation with finer-scale scheduling data.

Table 1. Integrated Benchmark Evaluation of Predictive Accuracy and Physical Constraint Feasibility on the 16-Month Test Set (June 2025 to September 2026) (a) Forecast accuracy and shock error.

Model Architecture	MAE (10^4 TEU)	RMSE (10^4 TEU)	MAPE (%)	R^2	Peak Shock Error (10^4 TEU)
ARIMA(1,1,1)	1.142	1.391	6.32%	0.479	2.254
SARIMA(1,1,1)(1,1,0)$_{12}$	0.856	1.084	4.73%	0.688	1.642
Standard BiLSTM	0.698	0.915	3.86%	0.781	1.388
Hybrid (w/o Shock Gate)	0.524	0.712	2.89%	0.871	0.975
Proposed (SARIMA–SG–BiLSTM)	**0.314**	**0.419**	**1.73%**	**0.956**	**0.538**

Table 2. Network feasibility and computation.

Model Architecture	Flow Balance Bias (%)	Capacity Violation (%)	Max Overflow (TEU)	Optimization Run-Time (s)
ARIMA(1,1,1)	4.85%	31.25%	12,460	0.42
SARIMA(1,1,1)(1,1,0)$_{12}$	3.12%	25.00%	8,720	0.45
Standard BiLSTM	2.45%	18.75%	6,580	1.18
Hybrid (w/o Shock Gate)	1.68%	12.50%	4,210	1.25
Proposed (SARIMA–SG–BiLSTM)	**0.42%**	**0.00%**	**0**	**1.36**

Note: Predictive metrics evaluate errors between forecasted demand $\hat{y}_{t+1}$ and empirical throughput y_{t+1}. Flow Balance Bias is calculated as $|\sum_k x_k - \hat{y}| / \hat{y}$. Capacity Violation Rate denotes the percentage of test months where unadjusted demand assignments exceed gateway physical capacity limits (U_k). Optimization Run-Time measures the mean wall-clock solving duration per monthly allocation cycle using the HiGHS solver.

4. Operational Insights, Policy Implications, and Limitations

4.1. Discussion of Forecast and Allocation Results

The test period includes changes in freight volume around the Red Sea shipping disruption and the May 2026 peak of 196,857 TEU. The comparison indicates that seasonal patterns alone did not explain the observed variation. It also shows why a demand forecast should be evaluated alongside the capacity of the corridors to which traffic is assigned.

SARIMA provides a seasonal baseline, while the gated BiLSTM adjusts its residuals as the

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

disruption indicator rises. The gate approaches a high weight ($g_t \rightarrow 1.0$) during the evaluated disruption period. The combined model reached 1.73% MAPE and a peak residual error of 5,380 TEU. The ablation supports a contribution from the gate within this test set; the gain under other disruptions remains untested.

The network results add a separate constraint to forecast accuracy. At the May 2026 peak, the static allocation placed Erenhot at 89.7% and Alashankou at 87.7% of their stated capacities. These values are close to the 90.0% threshold used in the analysis. The scenario model evaluates alternative allocations when demand or available capacity changes.

In the modeled allocation, freight is shifted toward Suifenhe and the Trans-Caspian Middle Corridor (TITR). Reported gateway utilization remains below 85%, the capacity violation rate falls to 0.00%, and total transport cost differs by 2.3%. These are model outcomes under the stated capacity and cost assumptions; implementation would require corridor-specific operational validation.

4.2. Policy and Practical Recommendations for CREX Governance

The results suggest three practical uses for the framework: monitoring external disruption indicators, comparing allocations across corridors, and reviewing cross-border handling constraints.

4.2.1. Disruption Monitoring and Capacity Planning

At the operational enterprise level, regional CREX coordination platforms (such as the China Railway Express Coordination Committee and provincial operating consortia) could use disruption indicators in monthly capacity planning. Coordination platforms could combine data linking customs tracking interfaces with real-world maritime intelligence, including Suez and Panama transit reports, the Shanghai Containerized Freight Index (SCFI), and international border clearance metrics. By incorporating the shock-gated mechanism developed in this study into enterprise resource planning (ERP) systems, dispatchers can inform capacity-buffering decisions whenever external risk indices surpass calibrated activation thresholds ($g_t \geq 0.50$). This capability could help operators review train paths and rolling-stock availability before expected demand increases.

4.2.2. Corridor Capacity Balancing

The modeled peak allocation uses spare capacity at Suifenhe and the TITR while reducing utilization at Erenhot and Alashankou. Dispatching platforms could test similar reallocations against actual train paths, ferry capacity, costs, and border procedures. Expansion of secondary

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

corridors would require separate demand and infrastructure assessments; the model alone does not establish an investment case.

4.2.3. Cross-Border Transshipment Coordination

Track-gauge differences require container transshipment at several border gateways. Electronic data interchange and mutual recognition of customs documentation could reduce administrative handling time, although those effects were not estimated in this study. Corridor pricing or slot allocation could also be evaluated as a way to manage peak demand, subject to regulatory and operational review.

4.3. Research Limitations and Future Directions

The analysis has three limitations:

(1) Data frequency: Monthly totals across six border ports do not resolve daily train assembly, yard queues, or short-term track closures. Daily or weekly records would allow those processes to be tested.

(2) Endogenous Capacity Reductions and Multimodal Competition: The optimization model in this study treats physical capacity degradation via an exogenous perturbation factor ($\lambda\delta_k$) and assumes static unit freight rates. In practice, shipping lines, air freight operators, and railway platforms adjust their freight tariffs dynamically in response to market demand, which may change mode choice and available capacity. Subsequent studies could integrate game-theoretic pricing models or bi-level programming formulations to capture the competitive dynamics between maritime carriers and rail operators under disruption regimes.

(3) Spatial features: The shock gate uses disruption indicators but does not explicitly represent how a disturbance propagates through rail junctions, dry ports, and intermodal terminals. Future work could test graph-based residual models against the present benchmark.

5. Conclusion

This study combines a SARIMA baseline, shock-gated BiLSTM residual correction, and scenario-based corridor allocation for CREX freight planning. On the reported 16-month test set, the forecast reached 1.73% MAPE and a coefficient of determination of 0.956. In the modeled peak-month allocation, gateway capacity violations were 0.00% and total cost varied by 2.3%. These findings support evaluating forecast accuracy together with corridor capacity. The monthly data and assumed costs and capacities limit direct transfer to day-to-day dispatching; further validation should use finer operational records and additional disruption periods.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

References

[1] Besharati, H., Nejad, M. S., & Govindan, K. (2022). Resilience and vulnerability in inter modal freight transport networks: A systematic review. Transportation Research Part E: Logistics and Transportation Review, 165, 102850.

[2] Zhao, Y., Cong, L., & Liang, X. (2023). Evolution and driving mechanisms of the China –Europe Railway Express network: A complex network perspective. Journal of Transpor t Geography, 106, 103522.

[3] Sun, X., Wandelt, S., & Zhang, A. (2022). Impact of geopolitical disruptions on Eurasia n railway logistics: Vulnerability, adaptability, and policy implications. Transport Policy, 129, 112–124.

[4] Hoffmann, J., Kumar, V., & Sanchez, M. (2024). The Red Sea shipping crisis and global trade disruption: Navigating rerouting costs and supply chain delays. Maritime Economi cs & Logistics, 26(2), 185–204.

[5] Yang, L., Xiao, Y., & Gu, W. (2024). Modal shift from maritime to rail freight under ge opolitical conflicts: An empirical investigation of the Eurasian land bridge. Transportatio n Research Part A: Policy and Practice, 182, 104018.

[6] Wang, C., Xu, Z., & Meng, Q. (2023). Bottleneck identification and transshipment capac ity enhancement for China-Europe container rail gateways. Transportation Research Part C: Emerging Technologies, 153, 104218.

[7] Notteboom, T., Pallis, P., & Rodrigue, J. (2023). Supply chain disruptions and container freight resilience in Eurasian corridors. Research in Transportation Business & Manage ment, 48, 100982.

[8] Lin, Z., Zhang, T., & Zhou, X. (2023). A critical review of multimodal freight demand f orecasting: Statistical, machine learning, and deep neural perspectives. Transport Revie ws, 43(5), 885–912.

[9] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.

[10] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learn ing forecasting methods: Concerns and ways forward. PLoS ONE, 13(3), e0194889.

[11] Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeser ies forecasting. Expert Systems with Applications, 37(1), 479–489.

[12] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. In Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Stru cture in Statistical Translation (pp. 103–111). Association for Computational Linguistics.

[13] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computatio n, 9(8), 1735–1780.

[14] Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Tr ansactions on Signal Processing, 45(11), 2673–2681.

[15] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

[16] Pai, P., & Lin, C. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33(6), 497–505.

[17] Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network m odel. Neurocomputing, 50, 159–175.

[18] Chen, Y., Luh, P. B., Guan, C., Zhao, Y., Michel, L. D., Coolbeth, M. A., Friedland, P. B., & Rourke, S. J. (2010). Short-term load forecasting: Similar day-based wavelet neura l networks. IEEE Transactions on Power Systems, 25(1), 322–330.

[19] Jiang, D., Fan, L., & Gao, Z. (2021). Hybrid modeling of railway passenger demand und er seasonal variations and holiday interventions. Journal of Advanced Transportation, 20

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.

21, 6652419.
[20] Song, H., Li, G., Witt, R., & Fei, B. (2019). Tourism demand modelling and forecasting: A review of recent research. Tourism Management, 75, 333–362.
[21] De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. Internati onal Journal of Forecasting, 22(3), 443–473.
[22] Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. Manageme nt Science, 66(6), 2445–2469.
[23] Elmachtoub, A., & Grigas, P. (2022). Smart "predict, then optimize". Management Scien ce, 68(1), 9–26.
[24] Savelsbergh, M., & Song, M. (2020). Inventory routing with stochastic demands: A revi ew. Computers & Operations Research, 113, 104803.
[25] Crainic, T., Perboli, G., & Rei, W. (2021). Scenario-based planning for multimodal freig ht transport operations under uncertainty. Transportation Research Part B: Methodologic al, 147, 22–45.
[26] Dong, X., He, Z., & Ke, Y. (2023). Capacity bottleneck mitigation at railway border che ckpoints: A transshipment scheduling approach. Transportation Research Part E: Logisti cs and Transportation Review, 171, 103038.
[27] Meng, L., Zhou, B., & Guo, X. (2023). Collaborative optimization of train scheduling an d container routing for the China-Europe railway express. Computers & Industrial Engin eering, 176, 108938.
[28] Besbes, B., Phillips, R., & Zeevi, A. (2022). Robust optimization for revenue manageme nt under demand learning and capacity constraints. Operations Research, 70(3), 1530–15 49.
[29] Wang, D., & Bayier, L. (2026). Forecasting Freight Flows and Mitigating Disruption Ris ks in the China–Europe Multimodal Transport Corridor: A Physics/Economics-Informed Spatiotemporal Diffusion Bi-LSTM Approach. International Journal of Advanced AI A pplications, 2(12), 60–76.
[30] Guo, D., Li, Z., & Tao, T. (2025). Bio-Inspired Adaptive Dynamic Attention: An Empiri cally Driven AI Framework for Human–Machine Coaching in Team Collaborative Decis ion-Making. International Journal of Advanced AI Applications, 1(8), 22–38.
[31] Wang, X., & Guo, D. (2025). Adaptive-Gated Spiking Neural Networks with Memristiv e Crossbars for Real-Time Athlete Injury Prediction. International Journal of Advanced AI Applications, 2(1), 1–16.

© 2027 The Author(s). Published by Hong Kong Dawn Clarity Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/.