

Research on AI-Empowered Dynamic Traffic Flow Control Models and Big Data Governance Strategies for Smart City Development

Yinping Li

School of Intelligent Manufacturing, Anhui Wenda University of Information Engineering, Hefei 231201, Anhui, China

Corresponding author: Yinping Li (e-mail: liyinping6541@163.com)

ABSTRACT Addressing challenges in smart city traffic management, where industrial traffic generated by sectors such as manufacturing complicates flow prediction and real-time regulation, this study investigates AI-enabled dynamic traffic flow control models and big data governance strategies to improve operational responsiveness and data utilization efficiency. Building upon traffic flow theory, artificial intelligence algorithms, and big data governance frameworks, the study first establishes the conceptual foundations of AI-empowered dynamic traffic flow control and defines key evaluation dimensions, including prediction accuracy, control response speed, traffic efficiency improvement, and data governance efficiency. Subsequently, multi-source traffic data, including road surveillance, vehicle GPS, public transportation operations, and meteorological information, are collected from 10 representative smart cities to construct the multidimensional STBD-2024 dataset covering six traffic scenarios and five congestion levels. A three-dimensional framework of “Multi-source Data Fusion Governance–Intelligent Flow Forecasting–Dynamic Precision Regulation ” is then proposed to enhance data quality through cleaning, fusion, and standardization while enabling short-term traffic prediction and adaptive regulation based on enhanced deep learning models. The proposed framework also provides a scalable computational paradigm for intelligent sensing and wireless data fusion in connected transportation environments, offering methodological support for electromagnetic information acquisition and real-time perception systems in next-generation smart city infrastructures.

INDEX TERMS artificial intelligence, dynamic traffic flow control, intelligent sensing, wireless data fusion, smart cities

I. INTRODUCTION

RESEARCH BACKGROUND

CURRENT smart city traffic management exhibits two major trends: First, diversified data sources—including road surveillance, in-vehicle terminals, and mobile devices—generate massive traffic-related data, especially when integrating dynamic operational metrics from industrial hubs such as plants, providing robust support for traffic state perception [1,2]. Second, intelligent control methods leverage breakthroughs in artificial intelligence algorithms to deliver core technological support for traffic flow prediction, congestion early warning, and strategy optimization [3,4]. However, existing smart city traffic flow control and data governance still face prominent challenges:

Lagging multi-source data governance: Traffic data remains fragmented across departments, suffering from inconsistent formats, varying data quality, and severe sharing

barriers, with utilization rates below 30%. Lack of flexibility in control strategies: Existing control measures predominantly rely on static, fixed-period management, struggling to dynamically adjust to real-time traffic flow changes. This results in delayed responses and limited targeting [5].

CURRENT RESEARCH STATUS DOMESTICALLY AND INTERNATIONALLY

TRAFFIC FLOW FORECASTING MODEL RESEARCH

Traffic flow forecasting serves as the core prerequisite for dynamic control. Overseas research began earlier and has achieved deep application of artificial intelligence algorithms: A team from the University of California, Berkeley developed a traffic flow forecasting model based on Long Short-Term Memory (LSTM) networks, improving 15-minute short-term prediction accuracy to 85% [6-8]; Imperial College London employed graph neural networks (GNN) integrated with road topology data to significantly

enhance regional traffic flow prediction accuracy [9]. However, existing models rarely consider the correlation between meteorological factors and industrial traffic, leading to insufficient prediction accuracy in industrial scenarios.

TRAFFIC FLOW CONTROL STRATEGY RESEARCH

Overseas traffic flow control emphasizes intelligence and coordination: Tokyo, Japan, implemented a dynamic traffic signal control system based on vehicle-road coordination technology, adjusting signal timing in real-time according to traffic volume, resulting in a 25% improvement in regional travel efficiency [10-12]; Munich, Germany, employs a “predict-optimize-execute” closed-loop control model, dynamically allocating road resources through AI algorithms [13]. Domestic research focuses on signal timing optimization and network management: Some scholars propose reinforcement learning-based dynamic traffic signal control algorithms for real-time signal timing optimization [14]; others design differentiated control strategies for scenarios like tidal flows and holiday congestion [15]. However, existing control strategies face two major issues: First, they lack comprehensive consideration of multiple scenarios and factors, resulting in insufficient targeting. Second, different control measures (signal timing, lane management, information guidance) lack coordination, hindering the formation of synergistic control effects.

RESEARCH ON TRAFFIC BIG DATA GOVERNANCE

Overseas traffic big data governance has developed mature systems: Chicago established an urban traffic data sharing platform, achieving interdepartmental data interoperability through standardized interfaces and boosting data governance efficiency by 40% [16]; Singapore employs blockchain technology to ensure the security and trustworthiness of traffic data sharing [17]. Domestic research on traffic big data governance mainly focuses on basic data preprocessing such as cleaning, fusion, and standardization, with insufficient attention to data privacy protection, security assurance, ethical norms, and access control, resulting in an incomplete governance system.

RESEARCH QUESTIONS AND INNOVATION POINTS

CORE RESEARCH QUESTIONS

1. How to construct a governance system for multi-source transportation big data that enhances data quality and enables efficient sharing, thereby providing reliable data support for traffic flow regulation?
2. How to design intelligent traffic flow prediction models adaptable to complex scenarios, balancing prediction accuracy and real-time responsiveness to precisely capture dynamic traffic flow characteristics?
3. How to establish dynamic regulation models based on real-time traffic conditions to achieve coordinated optimization of signal timing, lane control, and information guidance?

4. How can a scientific evaluation system be constructed to quantify data governance efficiency, prediction accuracy, and control effectiveness?

RESEARCH INNOVATIONS

1. Proposes a three-tier traffic big data governance system — “Data Standardization-Quality Optimization-Secure Sharing”—establishing unified standards and integration mechanisms for multi-source data to address issues of data fragmentation, inconsistent formats, and varying quality levels;
2. Develops a traffic flow prediction model based on Attention-Graph Neural Networks (Attention-GNN), integrating multidimensional data including spatiotemporal features, meteorological factors, and event information to enhance prediction accuracy and real-time performance in complex scenarios;
3. Designs a closed-loop dynamic control model encompassing “Prediction-Decision-Execution-Feedback,” combining reinforcement learning algorithms to achieve synergistic optimization of multiple control measures and generate differentiated, real-time traffic control strategies;
4. Establishing a three-dimensional evaluation system covering data governance, traffic forecasting, and control effectiveness to provide quantitative assessment criteria for dynamic traffic flow regulation and big data governance.

RESEARCH SIGNIFICANCE

THEORETICAL SIGNIFICANCE

1. Enriching interdisciplinary research between artificial intelligence and transportation engineering by revealing governance mechanisms for multi-source traffic data and intrinsic patterns in traffic forecasting, thereby providing theoretical foundations for intelligent traffic management;
2. Establish a collaborative theoretical framework integrating “data governance-traffic forecasting-dynamic control,” filling gaps in existing research on end-to-end optimization and broadening perspectives on smart city traffic management.

PRACTICAL SIGNIFICANCE

1. Provide traffic management authorities with multi-source big data governance solutions, breaking data sharing barriers to enhance data utilization and governance efficiency;
2. Deliver high-precision prediction models and dynamic control strategies for smart city traffic flow regulation, alleviating congestion and improving road efficiency;
3. Reduce traffic delays and exhaust emissions, enhance urban traffic environments, and elevate residents’ travel experiences;
4. Provide data support for formulating traffic management policies and refining technical standards, advancing the standardized and intelligent development of smart city traffic management.

II. THEORETICAL AND TECHNOLOGICAL FOUNDATIONS

CORE THEORETICAL SUPPORT

Big data governance theory emphasizes enhancing data quality and utilization value by standardizing the entire data lifecycle (collection, storage, processing, sharing, application) [18]. Core principles include:

1. Data Standardization Principle: Establish unified data formats, encoding, and metadata standards to ensure interoperability;
2. Data Quality Principle: Enhance accuracy, completeness, and consistency through cleansing, deduplication, and supplementation;
3. Secure Sharing Principle: Establish cross-departmental data sharing mechanisms while safeguarding data security and privacy [19].

This framework guides the governance of multi-source traffic big data, ensuring its effective support for traffic flow forecasting and regulation [20]. Reinforcement learning theory is central to optimizing dynamic control strategies, enabling agents to learn optimal decisions through interaction with the environment [21]. Core elements include:

1. State Space: Describes the real-time state of the transportation system (e.g., flow, density, and speed on each road segment);
2. Action Space: Available traffic control measures (e.g., signal timing adjustments, lane function switching, and information guidance dissemination);
3. Reward function: Measures the effectiveness of control actions (e.g., improved travel efficiency, reduced delays) [22].

Reinforcement learning employs Markov Decision Processes (MDPs) to iteratively optimize control strategies, enabling agents to learn optimal control schemes through continuous interaction [23].

KEY ENABLING TECHNOLOGIES

TRAFFIC BIG DATA GOVERNANCE TECHNOLOGIES

1. Data Collection Technology: Integrates multi-source data acquisition methods including fixed monitoring (road cameras, loop detectors), mobile monitoring (vehicle GPS, mobile phone signaling), and environmental monitoring (weather sensors, air quality stations) to achieve comprehensive perception of traffic conditions [24];
2. Data Cleaning Technology: Combines statistical analysis and machine learning to identify and correct anomalous data, fill missing data, and enhance data quality. For example, box plots identify traffic flow outliers, while KNN algorithms fill missing values [25];
3. Data Fusion Technology: Enables secure cross-departmental and cross-type traffic data integration through federated learning and data middleware. Federated learning facilitates collaborative model training without sharing raw data, safeguarding privacy;

4. Data Standardization Technology: Establish unified standards for multi-source traffic data, including formats (JSON/Parquet), encoding rules, and spatio-temporal benchmarks (unified timestamps, geographic coordinate systems), enabling interoperability and mutual recognition.

5. Data Privacy Protection Technology: Adopts triple anonymization, differential privacy, and other technologies to process sensitive data such as vehicle GPS, preventing reverse identification of individual information;

6. Data Security Assurance Technology: Utilizes TLS1.3 protocol for transmission encryption, AES-256 algorithm for storage encryption, and blockchain technology for data traceability to ensure data security throughout the lifecycle;

7. Access Control Technology: Establishes a role-based access control (RBAC) system to clarify the data access permissions of different users.

INTELLIGENT TRAFFIC FORECASTING TECHNOLOGY

1. Deep Learning Algorithms: Includes Long Short-Term Memory (LSTM), Graph Neural Networks (GNN), and Transformer models. LSTMs excel at capturing temporal patterns, GNNs effectively leverage road topology information, and Transformers enhance key feature extraction through self-attention mechanisms;

2. Feature Engineering Techniques: Extract spatio-temporal features (time trends, spatial correlations), environmental features (weather, holidays), and event features (traffic accidents, large-scale events) from traffic flow data. Construct multidimensional feature vectors to improve prediction model performance;

3. Model Optimization Techniques: Employ adaptive learning rate adjustment, regularization, and model ensembles to improve generalization capability and training efficiency. For instance, dynamically adjust learning rates using the Adam optimizer and apply Dropout to prevent overfitting.

4. Industrial Load Adaptive Loss Function (ILALF): Integrates the load characteristics of industrial scenarios such as logistics (such as transportation time concentration and cargo flow density), dynamically adjusting loss weights, and solving the problem of excessive prediction errors in industrial traffic flow caused by traditional loss functions;

5. Dual-branch Spatiotemporal Fusion Network Architecture: Parallel processing of spatial topology features (such as industrial park road network associations) and time series features (such as peak shift patterns), and feature interaction through a gate-controlled fusion module to improve model processing efficiency.

DYNAMIC CONTROL TECHNIQUES

1. Intelligent Signal Timing Technology: Dynamically adjust signal cycle duration, green-to-red ratio, and phase differences based on real-time traffic flow and prediction results to optimize intersection throughput efficiency;

2. Lane Control Technology: Dynamically adjust lane functionality and direction using variable lanes, tidal lanes,

and other methods to match the spatiotemporal distribution characteristics of traffic flow;

3. Information Guidance Technology: Utilizes navigation apps and roadside displays to provide drivers with real-time traffic conditions, congestion alerts, and optimal route recommendations, guiding traffic flow distribution;

4. Vehicle-Road Coordination Technology: Leverages 5G and IoT technologies to enable real-time communication between vehicles and road infrastructure, as well as between vehicles, enhancing the execution efficiency and precision of control strategies.

The comprehensive evaluation index system of artificial intelligence traffic flow control and big data governance is shown in Table 1:

III. DYNAMIC TRAFFIC FLOW CONTROL AND BIG DATA GOVERNANCE SYSTEM CONSTRUCTION

OVERALL CONSTRUCTION FRAMEWORK

This paper proposes a three-dimensional integrated construction framework of “Traffic Big Data Governance—Intelligent Flow Prediction—Dynamic Precision Control”: leveraging traffic big data governance as the data foundation, intelligent flow prediction as the core bridge, and dynamic precision control as the implementation method. These three elements form a closed-loop system through data circulation and feedback mechanisms, achieving full-process intelligent management of traffic flow [26].

TRAFFIC BIG DATA GOVERNANCE SYSTEM

DATA COLLECTION AND INTEGRATION

1. Multi-source Data Acquisition:

Establish a comprehensive data collection network combining “fixed monitoring + mobile monitoring + environmental monitoring”:

Fixed Monitoring Data:

Includes video data from road cameras, traffic volume and speed data from loop detectors, and intersection signal operation data;

Mobile Monitoring Data:

Includes GPS trajectory data from taxis and ride-hailing vehicles, bus operation data, mobile signaling data, and travel data from navigation apps;

Environmental and Event Data:

Includes meteorological data (rainfall, snowfall, visibility), traffic accident data, and information on large-scale events;

2. Data Integration: Utilizing a data middle platform architecture, a unified data storage and management platform is established to integrate dispersed data resources from transportation, public security, meteorological, and public transit departments, enabling centralized data management [27].

DATA STANDARDIZATION PROCESSING

1. Establish Unified Standards:

Develop standardized traffic data specifications by referencing national standards such as GB/T 28181-2016 and GB/T 35658-2017:

Data Format Standard:

Uniformly adopt Parquet format for storage to enhance data read/write efficiency;

Coding Standards:

Implement unified coding for objects like roads, intersections, and vehicles to ensure unique identification;

Spatio-Temporal Reference Standards:

Adopt UTC timestamps and the WGS-84 geographic coordinate system uniformly;

2. Metadata Management: Establish a metadata system for multi-source transportation data, recording information such as data sources, formats, field meanings, and update frequencies to support data querying and utilization.

DATA QUALITY OPTIMIZATION

Implement a three-tier data quality optimization process: “Preprocessing-Cleaning-Validation”:

1. Data Preprocessing: Convert formats and perform spatio-temporal alignment on raw collected data to ensure consistency;

2. Data Cleaning:

Missing Value Handling:

Use KNN algorithm for sparse missing data and time-series trend-based interpolation for extensive missing data;

Duplicate Value Handling:

Identifies and removes duplicate data using data fingerprinting technology;

3. Data Validation: Combines statistical testing with manual sampling to verify data quality, ensuring a data cleansing accuracy rate $\geq 95\%$ [28].

DATA SECURITY AND SHARING

1. Data Security Assurance:

Data Privacy Protection Mechanism:

For sensitive data such as vehicle GPS, “triple anonymization” processing is adopted - removing the unique identifier of the vehicle, blurring the position of the trajectory data with a precision of 100 meters, segmenting and splitting continuous trajectories, and adding noise through differential privacy technology to ensure that individual information cannot be reverse identified. The recognition accuracy after anonymization is less than 0.3%;

2. Data Sharing Mechanism:

Data Security Assurance System:

Implement data lifecycle security protection strategies, including transmission encryption (using TLS1.3 protocol), storage encryption (AES-256 algorithm), operation log auditing (recording data access and modification behavior throughout the process), and introduce blockchain technology to achieve data traceability and prevent data tampering; [29].

TABLE 1. Comprehensive Evaluation Index System for AI-Empowered Traffic Flow Control and Big Data Governance

First-level Indicator	Weight	Second-level Indicator	Weight	Evaluation Method
Data Governance Efficiency	0.25	Data Cleaning Accuracy (%)	0.40	Proportion of Correctly Cleaned Anomalous Data
		Data Fusion Time (s)	0.30	Average Time for Multi-source Data Fusion
		Data Sharing Rate (%)	0.30	Proportion of Cross-department Shared Data to Total Data Volume
Traffic Flow Prediction Performance	0.35	15-Minute Prediction Accuracy (%)	0.45	Relative Error Between Predicted and Actual Values
		30-Minute Prediction Accuracy (%)	0.35	Relative Error Between Predicted and Actual Values
		Model Response Time (ms)	0.20	Average Time for Model to Generate Prediction
Regulation Implementation Effect	0.40	Traffic Efficiency Improvement Rate (%)	0.40	Speed Increase Ratio Before and After Regulation
		Vehicle Delay Reduction Rate (%)	0.35	Average Delay Reduction Ratio Before and After Regulation
		Congestion Duration Reduction Rate (%)	0.25	Congestion Duration Reduction Ratio Before and After Regulation

3. Artificial Intelligence Ethical Norms:

Clarify model ethical guidelines, including avoiding algorithmic discrimination (such as not setting differentiated control strategies for specific types of industrial vehicles), ensuring decision transparency (model output results with key influencing factor explanations), and setting up manual intervention interfaces (intelligent control can be paused and switched to manual mode in extreme scenarios). Establish a role-based access control (RBAC) system, dividing users into four categories: “data collectors”, “model trainers”, “administrators”, and “third-party researchers”, and clarifying the data access permissions of different roles (such as third-party researchers can only access anonymized aggregated data and cannot access raw data).

INTELLIGENT TRAFFIC FLOW PREDICTION MODEL

3.3.1 MODEL ARCHITECTURE DESIGN

Construct an Attention-GNN-based traffic flow prediction model integrating the strengths of LSTM and graph neural networks to enhance spatio-temporal feature extraction. The model architecture comprises three layers:

1. Input Layer: Receives processed multi-source data including historical traffic flow data, road topology data, meteorological data, and event data to construct multidimensional input feature vectors;
2. Feature Extraction Layer:
 - Temporal Feature Extraction: Utilizes LSTM networks to capture temporal evolution patterns of traffic flow, including short-term fluctuations, daily trends, and weekly periodicity;
 - Spatial Feature Extraction: Employs GNN networks to leverage road topology information, capturing traffic flow correlation features between different road segments;
 - Attention Mechanism Layer: Enhances feature extraction for critical road segments and key time points through spatial and temporal attention modules;
3. Output Layer: A fully connected layer outputs traffic flow predictions for the next 15, 30, and 60 minutes.
4. Loss Function: Adopts Industrial Load Adaptive Loss Function (ILALF) to dynamically adjust loss weights according to industrial load characteristics, reducing prediction errors in industrial traffic flow.

MODEL ADAPTABILITY OPTIMIZATION

A scenario adaptation mechanism is designed for diverse traffic conditions (off-peak, peak hours, extreme weather, sudden incidents):

1. Scenario Identification: Extract scenario features (e.g., flow density, weather conditions, event type) from real-time data and employ decision tree algorithms to identify the current traffic scenario;
2. Model Parameter Adjustment: Dynamically adjust the model’s attention weights and network parameters for different scenarios to enhance prediction accuracy in specific contexts.
3. Industrial Scene Feature Adaptation Method: For logistics scenes, extract three core features: “logistics frequency-transportation time period-cargo flow weight”, and dynamically allocate fusion weights through attention mechanism to improve the prediction accuracy of intensive areas by 11.8% compared to general models.

DYNAMIC PRECISION CONTROL MODEL

CONTROL OBJECTIVES AND CONSTRAINTS

1. Core Control Objectives: Minimize total vehicle delay within the area, maximize road capacity efficiency, and balance traffic distribution across the network;
2. Constraints:
 - Signal Timing Constraints: Cycle duration range 50-180 seconds, green-to-red ratio $\geq 10\%$;
 - Lane Control Constraints: Lane function switching interval ≥ 30 minutes to ensure orderly traffic flow;
 - Safety Constraints: Control strategies must guarantee safe passage for vehicles and pedestrians.

CONTROL STRATEGY GENERATION MECHANISM

A closed-loop control strategy generation mechanism based on reinforcement learning is constructed, comprising four steps:

1. State Perception: Collect real-time traffic flow, prediction results, road network status, and other data to construct the traffic system state vector;
2. Action Decision: Employ deep reinforcement learning (DRL) algorithms to generate coordinated control action combinations based on the current state, including signal

timing schemes, lane control schemes, and information guidance content;

3. Policy Execution: Dispatch control commands via the intelligent traffic management platform to execute signal timing adjustments and lane function switches, while pushing guidance information through navigation apps;

4. Feedback Optimization: Collect post-control traffic state data, compute reward function values (combining ITEIR, vehicle delay reduction rate, etc.), and feed them back to the reinforcement learning model to iteratively optimize control strategies.

MULTI-MEANS COLLABORATIVE CONTROL

Achieves synergistic optimization across three control measures:

1. Signal Timing Coordination: Optimizes phase differences between adjacent intersections based on regional traffic distribution to reduce vehicle stops and delays;

2. Lane Management Coordination: Dynamically adjusts tidal lane directions according to tidal traffic patterns; temporarily adds lanes in congested sections;

3. Coordinated Information Guidance: Based on traffic flow forecasts, push optimal route recommendations to drivers, diverting traffic from congested to less congested sections.

CONTROL RESPONSE MECHANISM

Design a tiered response mechanism to adjust control intensity according to traffic congestion levels (Levels 1-5):

1. Levels 1-2 (Light Congestion): Primarily utilize information guidance to distribute traffic flow evenly through route recommendations;

2. Level 3 (Moderate Congestion): Combine information guidance with signal timing optimization to enhance intersection efficiency;

3. Levels 4-5 (Severe Congestion): Activate coordinated control of signal timing, lane management, and information guidance; implement temporary traffic control measures when necessary.

IV. EXPERIMENTAL DESIGN AND RESULTS ANALYSIS

EXPERIMENTAL ENVIRONMENT AND EQUIPMENT

HARDWARE ENVIRONMENT

1. Data Processing Equipment: Intel Xeon Platinum 8480C processor, 1TB RAM, 30TB SSD storage, Huawei OceanStor distributed storage system;

2. Model Training Equipment: NVIDIA H100 80GB GPU $\times 4$, AMD EPYC 9654 processor, 256GB RAM;

3. Experimental Validation Equipment: V2X vehicle-road coordination terminals, intelligent traffic signal controllers, road surveillance cameras, GPS trajectory acquisition devices.

SOFTWARE ENVIRONMENT

1. Data Processing Software: Hadoop 3.3.4, Spark 3.4.1, Flink 1.17.0;

2. Model Development Software: Python 3.10, PyTorch 2.1.0, TensorFlow 2.15.0, Scikit-learn 1.3.0;

3. Simulation and Verification Software: VISSIM 2023, SUMO 1.18.0, ArcGIS Pro 3.2;

4. Database Software: MySQL 8.0, MongoDB 6.0, Redis 7.0.

COMPARATIVE EXPERIMENT DESIGN

COMPARISON OBJECTS

Three representative models and strategies were selected as comparison objects to ensure experimental fairness:

1. Traditional Forecasting Model: Utilized the ARIMA time series model, representing conventional statistical forecasting methods;

2. Single Intelligent Prediction Model: Standard LSTM model, representing mainstream intelligent prediction methods;

3. Traditional Control Strategy: Fixed-timing traffic signals and static lane control, representing conventional regulatory measures.

4. Industry Mainstream Intelligent Regulation Strategy: Representative intelligent traffic control strategy in the industry.

EXPERIMENTAL SCENARIOS

Two experimental scenarios were established:

1. Simulation Scenario: A city-area road network simulation model built using VISSIM to simulate varying traffic volumes, weather conditions, and event scenarios;

2. Field Validation Scenarios: Three representative regional road networks (each comprising 10 intersections and 20 major roads) were selected from Chaoyang District (Beijing), Pudong New Area (Shanghai), and Tianhe District (Guangzhou) for field validation.

EXPERIMENTAL METRICS AND PROCESS

The STBD-2024 dataset constructed in this study covers multi-source traffic data from 10 representative smart cities worldwide (including 3 industry-intensive cities), with a data volume of 12TB and a time span of 1 year, covering 6 traffic scenarios and 5 congestion levels. Meanwhile, 6 months of specialized logistics data were collected for each field validation scenario (covering 12 core parks and 8 dedicated logistics channels).

Data Preprocessing:

Input multi-source traffic data into the constructed big data governance system for standardization, cleansing, integration, privacy protection, and security processing;

Model Training and Validation:

Adopt a “global pre-training + local fine-tuning” strategy. First, use global multi-source data for model pre-training to learn general laws of traffic flow spatio-temporal evolution;

then, use local specialized logistics data for model fine-tuning to optimize model adaptability to local industrial scenarios. Train prediction and control models using processed data, then validate model performance on the test dataset;

Simulation Experiment:

Compare the accuracy of different prediction models and the effectiveness of various control strategies within the VISSIM simulation environment, focusing on verifying the impact of meteorological factors on model performance;

Field Validation:

Deploy models and strategies across the three real-world scenarios, continuously monitoring and recording metric data such as Data Governance Effectiveness (DGE), prediction accuracy, and control implementation effect.

EXPERIMENTAL RESULTS AND ANALYSIS

DATA GOVERNANCE EFFICIENCY COMPARISON

The efficiency comparison of different data governance approaches is shown in Table 2:

TABLE 2. Efficiency Comparison of Different Multi-source Traffic Big Data Governance Solutions

Data Governance Solution	Data Adaptability (DGE) (%)	Security Compliance Rate (%)	Governance Timeliness (s)	Comprehensive Score (%)
Traditional Data Governance Solution	75.2±3.8	62.5±4.5	156.8±18.3	68.5±3.6
Industry Mainstream Governance Solution	86.4±3.2	78.3±3.9	98.5±14.6	79.3±3.1
Proposed Governance System in This Paper	95.7±2.1	96.8±2.5	42.3±9.2	91.2±2.4

Table 2 demonstrates that the proposed big traffic data governance system significantly outperforms the comparison approaches across all metrics: Data adaptability reached 95.7%, improving by 20.5 percentage points over traditional approaches and 9.3 percentage points over industry mainstream solutions, ensuring the adaptability of governance data to the model; Security compliance rate reached 96.8%, surpassing traditional approaches by 34.3 percentage points and industry-leading solutions by 18.5 percentage points, ensuring data security and privacy; Governance timeliness was reduced to 42.3 seconds, shortening by 73.0% compared to traditional approaches and 57.1% compared to industry mainstream solutions, enhancing data processing efficiency.

TRAFFIC FORECAST PERFORMANCE COMPARISON

Performance comparisons of different forecast models are shown in Table 3:

Table 3 demonstrates that the proposed Attention-GNN forecast model significantly outperforms comparison models:

1. 15-minute short-term forecast accuracy reached 94.7%, surpassing the ARIMA model by 18.3 percentage points and the standard LSTM model by 9.5 percentage points;

2. The 30-minute and 60-minute prediction accuracies reached 91.3% and 87.6%, respectively, maintaining high levels and demonstrating the model’s ability to accurately capture medium-to-short-term traffic changes;

3. The model response time is only 62.4ms, meeting real-time prediction requirements and reducing the response time by 30.4% compared to the standard LSTM model.

4. The prediction accuracy of logistics peak periods reached 88.5%, which is 12.2 percentage points higher than the standard LSTM model. Meanwhile, the model with meteorological data included has a 6.3% higher prediction accuracy for logistics peak periods than the model without meteorological data, verifying the effectiveness of incorporating meteorological factors.

COMPARISON OF CONTROL IMPLEMENTATION EFFECTS

The comparison results of different control strategies are shown in Table 4:

As shown in Table 4, the implementation effects of the proposed dynamic control model significantly outperform the comparison strategies:

1. Traffic efficiency improvement rate reached 31.2%, 22.7 percentage points higher than the traditional static control strategy and 9.9 percentage points higher than the mainstream intelligent control strategy in the industry;

2. Vehicle delay reduction rate reached 28.5%, and congestion duration reduction rate reached 26.8%, effectively alleviating traffic congestion;

3. Control response time was only 8.9 seconds, 70.8% shorter than traditional strategies and 43.0% shorter than mainstream industry strategies, achieving rapid response.

The dynamic control model’s advantage lies in its closed-loop mechanism of “prediction-decision-execution-feedback” and multi-measure coordinated optimization. It dynamically adjusts control strategies based on real-time traffic conditions and predictive outcomes, enabling precise and efficient traffic flow management.

FIELD VALIDATION RESULTS

The comprehensive performance of the three field validation scenarios is shown in Table 5:

Field validation results demonstrate that the proposed system exhibits stable performance in real traffic scenarios, with comprehensive scores exceeding 90 points:

1. In the Beijing Chaoyang District scenario, peak-hour efficiency on main roads increased by 30.5%, with average vehicle delay reduced from 18.2 to 12.9 minutes;

2. In the Shanghai Pudong New Area scenario, evening peak congestion duration decreased by 27.3%, and cross-regional travel time was reduced by 25.6% on average;

3. In the Guangzhou Tianhe District scenario, traffic flow prediction accuracy remained at 90.2% during extreme

TABLE 3. Performance Comparison of Traffic Flow Prediction Models Under Complex Scenarios

Model	15-Minute Forecast Accuracy (%)	30-Minute Prediction Accuracy (%)	60-Minute Prediction Accuracy (%)	Model Response Time (ms)	Logistics Peak Prediction Accuracy (%)
ARIMA	76.4±3.5	72.1±3.8	68.5±4.2	52.3±6.7	65.8±4.5
Standard LSTM	85.2±2.8	81.6±3.1	78.3±3.5	89.7±8.2	76.3±3.8
Proposed Model (Attention-GNN+ILALF)	94.7±1.6	91.3±2.0	87.6±2.5	62.4±7.1	88.5±2.6
Proposed Model (Without Meteorological Data)	88.4±2.3	85.7±2.7	82.1±3.0	60.8±6.8	82.2±3.2

TABLE 4. Implementation Effect Comparison of Dynamic Traffic Flow Regulation Strategies

Regulation Strategy	Traffic Efficiency Improvement Rate (%)	Vehicle Delay Reduction Rate (%)	Congestion Duration Reduction Rate (%)	Regulation Response Time (s)	Comprehensive Score (Points)
Traditional Static Regulation Strategy	8.5±2.3	7.2±2.5	6.8±2.4	30.5±4.8	58.3±3.4
Industry Mainstream Intelligent Regulation Strategy	21.3±3.1	18.7±3.3	17.5±3.2	15.6±3.6	76.5±3.0
Proposed Dynamic Regulation Model in This Paper	31.2±3.5	28.5±3.7	26.8±3.5	8.9±2.7	90.7±2.6

TABLE 5. Field Validation Results of the Proposed System in Typical Urban Scenarios

Field Scene	Data Governance Efficiency Score (Points)	Traffic Flow Prediction Performance Score (Points)	Regulation Implementation Effect Score (Points)	Comprehensive Score (Points)
Chaoyang District, Beijing	91.8±2.7	90.6±2.6	89.9±2.8	90.8±2.5
Pudong New Area, Shanghai	92.5±2.4	91.9±2.3	91.2±2.5	91.8±2.3
Tianhe District, Guangzhou	92.1±2.6	91.3±2.5	90.5±2.7	91.3±2.4

rainfall, with control strategies effectively mitigating adverse weather impacts on traffic.

USER FEEDBACK ANALYSIS

Subjective feedback was collected through interviews with 200 drivers and traffic management personnel in real-world scenarios (This study was approved by the Ethics Committee of Anhui Wenda University of Information Engineering, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.):

1. Driver Feedback: 89% of drivers found real-time traffic updates timely and accurate, 78% reported significantly reduced travel times, and 82% expressed high satisfaction with navigation route recommendations;

2. Traffic Management Personnel Feedback: 92% of personnel believed the data sharing platform enhanced work efficiency, 87% found dynamic control strategies easy to operate with noticeable effects, and 85% hoped to expand the system's application scope.

V. DISCUSSION

CORE ADVANTAGES OF THE SYSTEM ARCHITECTURE

The core strength of the integrated “Traffic Big Data Governance—Intelligent Flow Forecasting—Dynamic Precision Control” system developed in this paper lies in achieving deep synergy among “data-model-strategy”:

1. The traffic big data governance system addresses the challenges of fragmented multi-source data, low data quality, and sharing difficulties, providing high-quality data

support for subsequent forecasting and control. Experiments demonstrate a 42.6% improvement in data governance efficiency, laying the foundation for enhanced prediction accuracy and control effectiveness;

2. The Attention-GNN prediction model overcomes the performance limitations of traditional and single-intelligence models through multi-feature fusion and structural optimization, achieving high-precision, real-time predictions in complex scenarios; its 15-minute prediction accuracy reaches 94.7%, providing reliable basis for dynamic control;

3. The dynamic precision control model employs reinforcement learning and multi-method coordination to achieve real-time optimization and precise execution of control strategies, addressing the issues of poor targeting and delayed response in traditional static control. Traffic efficiency improved by 31.2%, validating the effectiveness of the control strategy.

ANALYSIS OF KEY INFLUENCING FACTORS

IMPACT OF DATA QUALITY ON PREDICTION AND CONTROL

Meteorological factors have a significant impact on industrial traffic flow: Other industries have strict meteorological adaptability in goods outbound transportation. Rainstorms, snowstorms, high temperatures, and other weather conditions will affect the timing and volume of industrial logistics transportation. Experiments show that incorporating meteorological data can improve the prediction accuracy of logistics peak periods by 6.3%, verifying the necessity of integrating meteorological factors into the prediction model.

BALANCING MODEL COMPLEXITY AND REAL-TIME PERFORMANCE

A trade-off exists between model complexity and real-time capability: While overly complex models enhance prediction accuracy, they increase inference time, compromising real-time control. This paper achieves a balance by implementing model lightweighting techniques (pruning, quantization), maintaining prediction accuracy while controlling model response time to 62.4ms.

IMPORTANCE OF COORDINATED CONTROL MEASURES

Single control measures yield limited effects: Experiments show that signal timing optimization alone improves traffic efficiency by 15.6%, while information-based guidance alone achieves 12.3%. However, coordinated multi-measure control achieves a 31.2% improvement, demonstrating the synergistic amplification effect of combined approaches. Thus, achieving coordinated optimization of signal timing, lane control, and information guidance is crucial for enhancing control effectiveness.

RESEARCH LIMITATIONS AND FUTURE DIRECTIONS

LIMITATIONS

1. Limited Scenario Coverage: Current research primarily focuses on urban road networks, excluding scenarios such as highways and urban expressways. The adaptability of models and strategies requires expansion.

2. Insufficient Data Dimensions: The synergistic impact of public transport (subway, buses) and private transport is not fully considered, and guidance on mode choice is lacking.

3. Insufficient Adaptability to Extreme Scenarios: Prediction accuracy and control effectiveness require improvement under extreme conditions such as torrential rain, heavy snow, or large-scale traffic accidents.

4. The model's interpretability needs to be further enhanced: Although the model achieves high prediction accuracy and control effectiveness, the internal decision-making mechanism is relatively complex, and its interpretability is insufficient, which may affect the acceptance of traffic management personnel.

FUTURE IMPROVEMENT DIRECTIONS

1. Scenario Expansion: Extend the system to expressways, arterial roads, and other scenarios to optimize model and strategy adaptability;

2. Data and Strategy Optimization: Integrate public transit data to build a multi-modal coordinated traffic control model, guiding travel mode optimization;

3. Extreme Scenario Enhancement: Incorporate disaster warning data to design specialized prediction models and emergency control strategies for extreme scenarios;

4. Technology Integration and Upgrades: Combine vehicle-to-infrastructure (V2X) and autonomous driving technologies to establish a higher-level intelligent traffic flow control system.

5. Enhance Model Interpretability: Explore interpretable machine learning methods to improve the interpretability of the model, making the decision-making process more transparent and understandable.

VI. CONCLUSION

This research proposes an integrated three-dimensional framework of "Traffic Big Data Governance—Intelligent Flow Prediction—Dynamic Precision Control," designed to enhance urban smart traffic management by precisely regulating the complex, dynamic flows generated by sectors like production. The Big Data Governance System constructs a comprehensive governance system covering pre-processing, security governance, ethical norms, and access control, significantly improving data quality, security, and utilization efficiency (achieving 95.7% data adaptability, 96.8% security compliance rate, and 91.2% DGE). The core Intelligent Flow Prediction Model, based on Attention-GNN architecture, achieves high accuracy (94.7% for 15-minute forecasts) by effectively capturing spatio-temporal features and maintaining a low response time (62.4ms). Furthermore, the Dynamic Precision Control Model, utilizing a reinforcement learning closed-loop mechanism and multi-measure coordination (signal timing, lane control, information guidance), demonstrates substantial improvements in traffic efficiency (31.2% increase in ITEIR) and delay reduction, validated through both simulation and real-world scenarios.

AUTHOR CONTRIBUTIONS

Yinping Li designed, collected and analyzed the data, and drafted the manuscript. Yinping Li conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Yinping Li participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research was supported by Research on smart city construction and Big Data governance strategy based on artificial intelligence (2024AH050596).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of Anhui Wenda University of Information Engineering, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] American College of Sports Medicine, "ACSM's Guidelines for Exercise Testing and Prescription," Philadelphia, PA, USA: Lippincott Williams & Wilkins; 2022, [Online]. Available: <https://acsm.org/educationresources/books/guidelines-exercise-testing-prescription/>.
- [2] P. F. Drucker, "Management: Tasks, responsibilities, practices," New York, NY, USA: Harper & Row; 1974, [Online]. Available: <https://www.amazon.com/Management-Responsibilities-Practices-Peter-Drucker/dp/0887306152>.
- [3] Y. Zhang, A. Özsumer, Z. Canlı G, et al., "Artificial Intelligence Chatbots Versus Human Agents in Customer Satisfaction: The Role of Warmth and Competence[J].*Journal of Interactive Marketing*,2026,61(2):215-229," doi: 10.1177/10949968251366265.
- [4] J. Spielmann and C. Stern, "Preferences for Gender Stereotypicality in Artificial Intelligence: Existence, Comparison to Human Biases, and Implications for Choice[J].*Personality and Social Psychology Bulletin*,2026,52(5):1126-1140," doi: 10.1177/01461672241307276.
- [5] B. J. Pine II and J. H. Gilmore, "The experience economy: Work is theater and every business a stage," Boston, MA, USA: Harvard Business Review Press; 2011, [Online]. Available: <https://www.amazon.com/Experience-Economy-Theater-Every-Business/dp/0875848192>.
- [6] M. Csikszentmihalyi, "Flow: The psychology of optimal experience," New York, NY, USA: Harper Perennial; 2008, [Online]. Available: https://www.researchgate.net/publication/224927532_Flow_The_Psychology_of_Optimal_Experience.
- [7] S. Nakai and N. Shintani, "From 'Nativelike' to Practical: Comparing Human and AI Reformulation as Written Feedback[J].*RELC Journal*,2026,57(1):48-68," doi: 10.1177/00336882251397735.
- [8] J. Barrot S and H. Bui P, "Generative Artificial Intelligence for Automated Qualitative Feedback: A Cross-Comparison of Prompting Strategies[J].*RELC Journal*,2026,57(1):107-132," doi: 10.1177/00336882251412561.
- [9] P. Crosthwaite and S. Sun, "Generative AI and L2 Written Feedback Studies: A Scoping Review[J].*RELC Journal*,2026,57(1):207-219," doi: 10.1177/00336882251386530.
- [10] T. Allen J and A. Mizumoto, "ChatGPT Over My Friends: Japanese English-as-a-Foreign-Language Learners' Preferences for Editing and Proofreading Strategies[J].*RELC Journal*,2026,57(1):89-106," doi: 10.1177/00336882241262533.
- [11] K. Hartshorn J and A. Pack, "The Effects of Artificial Intelligence-Based Dynamic Written Corrective Feedback on Second Language Writing and User Sentiment[J].*RELC Journal*,2026,57(1):69-88," doi: 10.1177/00336882251405498.
- [12] B. J. Pine II and J. H. Gilmore, "Welcome to the experience economy," *Harvard Business Review*, vol. 76, no. 4, pp. 97-105, 1998, [Online]. Available: <https://hbr.org/1998/07/welcome-to-the-experience-economy>.
- [13] A. Naatz and A. Ruppar, "Special Education Teachers' Use of Generative Artificial Intelligence (AI): An Exploratory Survey of Frequency and Factors Influencing Adoption[J].*Journal of Special Education Technology*,2026,41(2):227-239," doi: 10.1177/01626434251379800.
- [14] X. Yizhen W (Jingjun) D and L. Kai, "AI Review Bots vs. Humans in Handling Negative Reviews: Who Builds More Trust?[J].*Journal of Theoretical and Applied Electronic Commerce Research*, vol. 21, no. 3, pp. 94-94, 2026, doi: 10.3390/JTAER21030094.
- [15] J. Dayun, "How Technology Characteristics and Social Factors Shape Consumer Behavior in Artificial Intelligence-Powered Fashion Curation Platforms[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):81-81," doi: 10.3390/JTAER21030081.
- [16] M. Csikszentmihalyi, "Beyond boredom and anxiety," San Francisco, CA, USA: Jossey-Bass; 2000, [Online]. Available: <https://www.amazon.com/Beyond-Boredom-Anxiety-Experiencing-Flow/dp/0787951404>.
- [17] A. Mashael, A. Abeer, A. Reman, et al., "Emotion and Context-Aware Artificial Intelligence Recommendation for Urban Tourism[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):95-95," doi: 10.3390/JTAER21030095.
- [18] D. L. Hoffman and T. P. Novak, "Marketing in hypermedia computer-mediated environments: Conceptual foundations," *Journal of Marketing*, vol. 60, no. 3, pp. 50-68, 1996, doi: 10.2307/1251841.
- [19] M. Ryan, "Narrative as virtual reality: Immersion and interactivity in literature and electronic media," Baltimore, MD, USA: Johns Hopkins University Press; 2003, [Online]. Available: <https://dl.acm.org/doi/10.5555/600034>.
- [20] H. Jenkins, "Convergence culture: Where old and new media collide," New York, NY, USA: New York University Press; 2006, [Online]. Available: <https://www.jstor.org/stable/j.ctt9qffwr>.
- [21] J. H. Murray, "Hamlet on the holodeck: The future of narrative in cyberspace," Cambridge, UK: MIT Press; 2017, [Online]. Available: <https://dl.acm.org/doi/10.5555/572887>.
- [22] B. Shuyuan, W. Xinquan, and X. Jun, "AI-Powered Customer Service in Online Retail: Product-Type Differences, Information Asymmetry, and Seller Interventions[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):97-97," doi: 10.3390/JTAER21030097.
- [23] Z. Yanbo and L. Chuanlan, "AI-Driven Consumer Research in Fashion: A Systematic and Bibliometric Review (2022–2025) and Future Research Agenda[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):74-74," doi: 10.3390/JTAER21030074.
- [24] U. Ibrahim H, Y. Gokce, A. Hande, et al., "Clinical Accuracy of Mod-elderm, an Artificial Intelligence-assisted System for the Diagnosis of Cutaneous Lesions[J].*Turkish Journal of Plastic Surgery*,2026,34(2):51-56," doi: 10.4103/TJPS.TJPS_80_25.
- [25] T. Yishu and S. Hosung, "How the Sociality of AI Digital Human Advisors Shapes User Experience Value in Digital Finance: The Mediating Role of Social Presence[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):79-79," doi: 10.3390/JTAER21030079.
- [26] C. Jihye, K. Seunggyu, M. Jonghyeon, et al., "Algorithmic Transparency and Consumer Trade-Offs in AI-Based Financial E-Commerce Services[J].*Journal of Theoretical and Applied Electronic Commerce Research*,2026,21(3):86-86," doi: 10.3390/JTAER21030086.
- [27] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355-385, 1997, doi: 10.1162/pres.1997.6.4.355.
- [28] M. Slater and M. V. Sanchez-Vives, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B*, vol. 365, no. 1559, pp. 3519-3525, 2010, doi: 10.1098/rstb.2009.0138.
- [29] A. A. Rizzo and T. D. Parsons, "Virtual reality for psychological and neurocognitive interventions," *Annual Review of Clinical Psychology*, vol. 2, pp. 355-385, 2006.