Design and Implementation of a Big Data Decision Support System for Balanced Allocation of Preschool Education Resources

Jiaxin Chen[1], Yang Yang[1], Wensheng Yan[2]
[1]Xi'an Fanyi University, Xi'an 710105, Shaanxi, China
[2]Quanzhou University of Information Engineering, Quanzhou 362000, Fujian, China
Corresponding author: Jiaxin Chen (e-mail: 15041531711@163.com)

ABSTRACT To address the dynamic imbalance between preschool education demand and resource allocation under rapid urbanization, this study proposes a spatiotemporal big data decision support system integrating multi-source data fusion, demand prediction, spatial accessibility assessment, and multi-objective optimization. The system adopts a front-end/back-end separation architecture and combines a Spatio-Temporal Graph Convolutional Network (STGCN), an improved Two-Step Floating Catchment Area (i2SFCA) method, and Multi-Objective Particle Swarm Optimization (MOPSO) to establish an automated workflow for demand forecasting, mismatch diagnosis, and resource configuration. WebGL- and WebGIS-based visualization modules enable interactive analysis and real-time decision support for large-scale spatial data. Experimental results on real urban datasets demonstrate that the proposed framework reduces the demand forecasting mean absolute percentage error to 6.8%, decreases the global spatial accessibility Gini coefficient to 0.28, and improves the supply-demand matching rate by 23% under the optimized allocation strategy. The proposed architecture provides an efficient engineering solution for heterogeneous information processing and large-scale spatial optimization, offering methodological references for intelligent sensing networks, distributed information integration, and data-driven decision support in electromagnetic and communication-related infrastructure systems.

INDEX TERMS spatiotemporal big data, decision support system, multi-source data fusion, spatial accessibility optimization, STGCN, MOPSO, electromagnetic information infrastructure

I. INTRODUCTION

In recent years, the accelerated pace of urbanization and the expansion of industrial clusters have triggered significant workforce migrations, causing the spatial and temporal distribution of the urban preschool-aged population to exhibit dramatic dynamic fluctuations [1-2]. As the starting point of the basic education system, the rationality of the spatial layout of preschool education directly relates to the realization of equal access to basic public services [3-4]. However, the current educational resource allocation models in most cities still heavily rely on static registered population bases and the traditional "per-thousand-person quota" allocation standard [5-6]. This experience-based static planning model, based on cross-sectional data, struggles to quickly capture the migration and spillover effects of micro-populations within complex urban systems. This results in the spatial planning of educational facilities lagging significantly behind the evolution of actual needs, frequently leading to a structural supply-demand mismatch where localized areas face difficulties in accessing preschool education while some newly built preschools have "unfilled places" [7-8]. The core solution to this supply-demand mismatch dilemma lies in the accurate analysis and quantitative constraints of key objects within the preschool education resource allocation system. The specific issues addressed in this paper mainly cover the dynamic evolution of the school-age population at the micro-block scale (demand-side issues) [9-10], the spatial distribution and assessed carrying capacity of existing and potential educational facilities (supply-side issues) [11-12], and the complex urban transportation network topology connecting the supply and demand sides (spatial impedance issues) [13-14]. In a dynamically fluctuating urban environment, the ability to proactively predict the evolution of the demand scale of school-age children within micro-spatial units and scientifically measure the matching degree of school places and spatial equity under real travel constraints is a necessary prerequisite for driving the transformation of traditional static planning to an agile and precise data-driven allocation paradigm [15-16].

Domestic and international scholars have conducted extensive research on the two specific aspects of micro-level population demand forecasting and public service facility site selection optimization. At the demand forecasting level, existing studies mostly use the Autoregressive Integrated Moving Average (ARIMA) model [17-18] or the grey prediction model for time-series extrapolation [19]. However, when dealing with high-density micro-grid data, they often ignore the spatial spillover effect of population caused by the development of adjacent plots, and face the limitations of prediction lag and periodic abrupt changes. At the supply and demand assessment and spatial optimization level, the traditional Two-Step Floating Catchment Area Method (2SFCA) [20-21] and gravity models are often used to measure spatial accessibility [22-23]. However, they are mostly based on idealized Euclidean straight-line distances and fail to fully incorporate the real traffic network impedance. At the same time, in the subsequent facility site selection decision, conventional heuristic algorithms often focus on minimizing construction costs or maximizing coverage, lacking a multi-objective trade-off mechanism for the two conflicting goals of "improving local spatial fairness" and "controlling overall travel efficiency".

To completely break down data barriers and overcome the static limitations of existing theoretical models, this research aims to design and implement a spatiotemporal big data decision support system for the balanced allocation of preschool education resources. Unlike traditional auxiliary tools that rely on single GIS (Geographic Information System) software for offline static planning, this system achieves deep innovation and a closed-loop architecture. At the system architecture and implementation level, a cross-departmental, multi-source, heterogeneous data fusion foundation was first constructed, abstracting the physical space into a standardized grid engine. Secondly, relying on a microservice architecture, core computing power such as Spatio-Temporal Graph Convolutional Network (STGCN), improved Two-Step Floating Catchment Area Method (i2SFCA), and Multi-Objective Particle Swarm Optimization (MOPSO) were containerized, constructing a fully automated data flow computing layer for "dynamic demand prediction—mismatch blind spot diagnosis—site selection and degree allocation." Finally, the system broke away from the traditional black-box decision-making model, developing a highly interactive visual auxiliary decision-making sandbox based on WebGL technology. This seamlessly transformed abstract mathematical calculations into intuitive spatio-temporal heatmaps and resource allocation indicator dashboards, completing the full engineering deployment from multimodal data access to visual decision-making solution output.

This project achieved the successful transformation of complex macroscopic theoretical optimal solutions into automated inference tools with micro-scale practical feasibility through the closed loop operation of the decision support system (DSS). Experimental and application results demonstrate the DSS's overall performance based on real-world urban datasets: at the prediction phase, the DSS reduced the mean accurate error percentage (MAPE) of street level predictions of the school aged population to 6.8% which was a significant improvement over conventional time series models; and at the configuration phase, the DSS's automated local site selection and capacity expansion schemes complied strictly with macro level fiscal construction constraints therefore reducing the Gini coefficient of preschool educational space accessibility to 0.28 and increasing the degree of school place supply and demand matching by 23%. This system not only achieves cross-disciplinary integration of spatiotemporal deep learning and heuristic spatial optimization within its underlying algorithm architecture, but also synchronizes regional education resource allocation with the dynamic demographic shifts triggered by manufacturing hub expansions to facilitate a system-level transition from traditional 'experience-based governance' to 'digital intelligent governance'.

II. CORE ALGORITHMS AND SYSTEM ARCHITECTURE FOR BALANCED RESOURCE ALLOCATION

CONSTRUCTION OF MULTI-SOURCE SPATIOTEMPORAL DATA FOUNDATION AND OVERALL SYSTEM ARCHITECTURE

To address the physical challenges of multi-source heterogeneous data barriers and inconsistent spatial granularity, the system's underlying layer constructs a spatiotemporal fusion engine based on gridded spatial mapping. The entire big data decision support system adopts a front-end and back-end separation paradigm and deeply integrates microservice-based GIS components. Its overall architecture, as shown in Figure 1, mainly consists of four core modules: multi-source data, persistent data layer, model computation layer, and front-end interaction layer.

Note: POI = Point of Interest, SQL = Structured Query Language, WMS = Web Map Service, WFS = Web Feature Service.

The collection and fusion of multi-source data forms the computational basis of the system. In specific engineering implementation, micro-level population data, road network data, and kindergarten points of interest are collected first. Then, a unified geospatial reference benchmark is established, and polygon overlay analysis and spatial interpolation algorithms are used to uniformly resample discrete attribute points and linear transportation networks, mapping them to discrete spatial grid cells of standard resolution. To support subsequent mathematical modeling and matrix operations, the urban study area is discretized into a set $G = \{g_1, g_2, \ldots, g_N\}$ containing N basic spatial cells. At any specific time stamp t, the multidimensional spatiotemporal attribute features fused by grid cell g_i are normalized and defined as feature vectors:

$$X_i^t = [P_i^t, S_i^t, C_i^t] \tag{1}$$

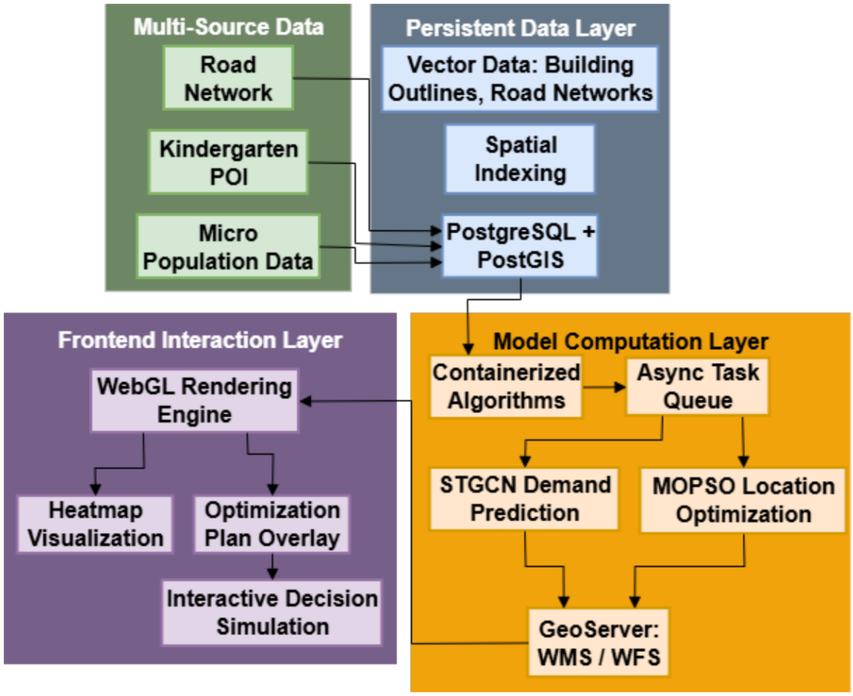

FIGURE 1. Overall framework diagram of the big data decision support system

In the formula, P_i^t represents the base number of the eligible population projected onto the grid after cleaning; S_i^t represents the equivalent supply of approved preschool education places covered by the current grid; and C_i^t is the connectivity feature of traffic nodes extracted based on complex networks. These operations resolve inconsistencies in data scale and structure, yielding a standardized computational matrix. For efficient access & storage of the spatial data, the persistent data layer is dependent upon a PostgreSQL database with a PostGIS extension that includes a spatial indexing method for quickly performing queries of the vector data that contain topological relationships (e.g., building footprints, street centre lines) and large combinations of features [24-25].

The model computation layer is at the heart of the computation capabilities of the system for decision making. Complex algorithms are contained within this layer and scheduled flexibly to utilize compute resources asynchronously via task queues for concurrent solution to the STGCN forecasting and MOPSO site selection optimization models. The computational output from the model is then provided to the GeoServer middleware where it is dynamically transformed into standard map tile services (WMS/WFS) [26-27].

The front-end interface allows terminal decision makers to use a visual representation of the interactivity of their terminal interface through its use of the WebGL rendering engine to allow users to visualise the spatial slices of a large resource of distributed heat maps and it integrates a live warning floating window to indicate i2SFCA mismatch blind spots. It will dynamically update so that users can compare the real-time position of the Pareto front for MOPSO optimisation with how users have interacted with the interface. The design of this system separates the heterogeneous data processing, the high-concurrent model solving, and the visualisation output from the physical medium, thus providing the user with a low-latency interactive simulation of decision making by providing a closed loop from the underlying data to the front-end business dashboard.

DYNAMIC PREDICTION OF MICRO-LEVEL PRESCHOOL EDUCATION DEMAND BASED ON STGCN

Demand for preschool education has changed through history because of how many babies were born each year, therefore an area will feel the effects over time depending on how many new homes and businesses were built nearby. The new system developed to analyze the patterns of preschool education (through graph convolutional networks) will take the grid based data sets for population flow and break them down to the correct times and locations of all of the previous years, in order to get appropriate data for future predictions [28-29]. To refine the demand estimation beyond simple resident density, the system introduces a choice-based demand weight coefficient (W_{choice}) into the STGCN input layer. This coefficient integrates multi-source proxy data, including school quality ratings and the spatial density of industrial work-places. By calculating the gravity-based correlation between residential grids and workplace hubs, the model acknowledges that parental school choice is a trade-off between home proximity, workplace convenience, and educational quality, rather than a naive mapping of residential population to the nearest facility. This nuanced approach ensures that the 6.8% MAPE reflects actual enrollment tendencies rather than just demographic presence. First, the study area is abstracted into a topological graph structure, where nodes represent basic spatial grids and edges reflect the physical spatial adjacency between grids. Historical grid population base sequences, cleaned from the base, are extracted. A time window length of τ is set, and cross-sectional data from consecutive time steps are superimposed to form an initial input tensor containing multi-dimensional spatiotemporal features.

To quantify the topological dependence of spatial migration of school-age children at the block scale due to the expansion of urban functional areas, the model uses graph convolutional layers to process spatial dimension information. By aggregating the attribute features of the target grid and its first-order neighboring nodes, the spatial agglomeration effect and population spillover characteristics are extracted. The layer-by-layer forward propagation calculation process of the graph convolution operator is defined as follows:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{2}$$

In the formula, l represents the hidden layer index of the graph convolutional network; $H^{(l)}$ is the input feature matrix of the l-th layer; $H^{(l+1)}$ represents the output feature matrix of the l-th layer after spatial aggregation; $\tilde{A}$ is the spatial adjacency matrix that introduces self-loops to preserve the self-attributes of grid nodes; $\tilde{D}$ is the degree matrix corresponding to matrix $\tilde{A}$; $W^{(l)}$ is the learnable weight parameter matrix of the l-th layer, used for linear transformation of the spatial feature dimension; σ is a non-linear activation function used to enhance the expressive power of the model.

Based on the completion of spatial dependency feature encoding, the model uses a one-dimensional temporal convolution structure to extract the dynamic laws of the sequence, addressing the non-linear evolution in the temporal dimension caused by population mobility and fertility rate fluctuations. A one-dimensional convolution operation is performed on the output of the spatial convolution along the time axis, and a gating mechanism is used to filter out irrelevant historical noise interference. The feature extraction equation in the temporal dimension is expressed as:

$$Y = \Gamma(H^{(l+1)} * K + b) \tag{3}$$

Y represents the spatiotemporal fusion feature matrix output by the temporal convolution module; Γ is the gated linear unit activation function, responsible for controlling the proportion of historical information transferred to future moments; $*$ identifies the standard one-dimensional convolution operator; K represents the convolution kernel parameters in the temporal dimension; and b represents the network bias term vector. Compared to traditional recurrent neural networks, this network structure effectively avoids the risk of gradient fading when processing long-period, high-density grid population data.

After alternating feature extraction in the spatial and temporal layers, the spatiotemporal feature tensor is flattened and input into the final fully connected layer for dimensionality mapping, ultimately generating the target prediction value. The mapping process is as follows:

$$\hat{P}_{t+1} = \Phi(Y) \tag{4}$$

In the formula, $\hat{P}_{t+1}$ is the matrix representing the predicted number of eligible children in each spatial grid cell at time $t+1$ in the future; Φ is the linear mapping function of the fully connected layer. Based on the above alternating mechanism of graph convolution and temporal convolution, the system transforms complex

historical trajectories into grid-scale demand predictions. To bridge the gap between quasi-real-time prediction and the 1-3 year construction cycle, the optimization module adopts a "rolling planning" strategy, using the 36-month long-term demand forecast as the configuration baseline rather than immediate snapshots.

As shown in Figure 2, the spatial features output by the STGCN model can intuitively and accurately map the evolution trajectory of the population center of gravity of school-age children over time. Within the continuously observed time steps (from the initial stage t1 to the predicted endpoint t12), the population center of gravity exhibits a unidirectional physical spatial migration from the highly saturated "old urban area" to the "new technology development zone." The model keenly captures the non-linear dynamic characteristics of this spatial transfer process, especially the "overflow transfer" inflection point accurately identified at time t7. Furthermore, as time progresses, the area of spatial nodes representing population density shows a stepped expansion trend, further verifying the algorithm module's extremely high fitting degree and forward-looking nature in handling micro-level population dynamic fluctuations and quantifying the spatial overflow effect of adjacent grids.

At the system engineering implementation level, the aforementioned STGCN prediction module has been extracted and encapsulated as an independent microservice component. This module is deployed in a computing engine running the PyTorch 2.0 framework, receiving time window parameters and prediction step sizes specified by the front end through a RESTful API (Representational State Transfer Application Programming Interface). In actual operation, the module periodically extracts cleaned spatiotemporal grid feature tensors from the persistent data layer for inference calculations, and writes back the output demand evolution indicators at each future time node to the data platform, directly serving as prior data input for the downstream supply and demand assessment and configuration modules.

SUPPLY AND DEMAND ASSESSMENT BASED ON AN IMPROVED TWO-STEP MOVING SEARCH METHOD

To accurately measure the supply-demand gap of existing preschool education resources under dynamic demand con-

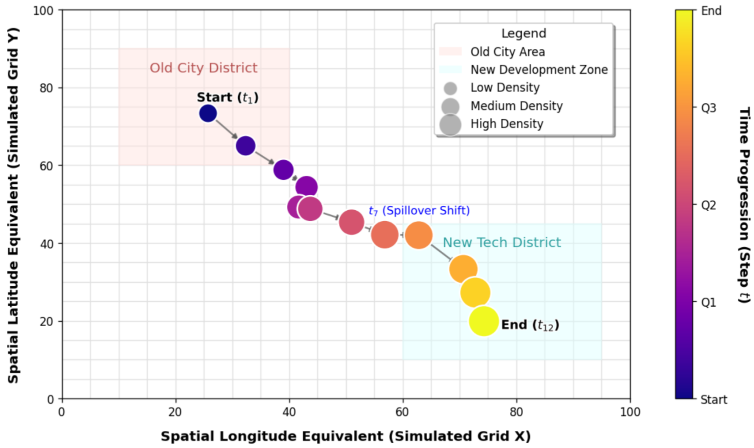

FIGURE 2. Spatial shift trajectory of the center of gravity of the school-age child population

straints, the system introduces an improved two-step moving search algorithm (i2SFCA) that considers the actual road network impedance [30-31]. The first step of the algorithm uses the actual carrying capacity of each kindergarten as the supply point and obtains all demand points within the extreme travel time threshold based on GIS road network topology analysis to quantify the actual carrying pressure on the supply side. The supply point is set as j, and its approved capacity (number of places) is S_j; the demand point is k, and its input value is the previously predicted number of school-age children, D_k. The mathematical process for calculating the supply-demand ratio R_j of supply point j is as follows:

$$R_j = \frac{S_j}{\sum_{k\in\Omega_j} D_k f(t_{kj})} \tag{5}$$

In the formula, t_{kj} is the shortest travel time from demand point k to supply point j along the actual road network; t_0 is the multi-modal travel time threshold matrix defined by user groups. While a 15-minute walking limit is used as the baseline for spatial equity, the system incorporates a multi-modal transport sensitivity analysis. This allows the threshold to dynamically adjust based on transport modes common in high-density areas, such as e-bikes ($t_0 = 8$ min) or public transit ($t_0 = 20$ min). Furthermore, the model applies a weather-safety impedance multiplier (β) to the road network topology, reflecting the increased physical limitations and reduced travel speeds during adverse conditions or in areas with lower urban safety scores, ensuring the allocation remains robust under varying common-sense constraints. $\Omega_j = \{k \mid t_{kj} \leq t_0\}$ represents the set of all demand points falling within the service area of supply point j; $f(t_{kj})$ is the distance decay function, and the denominator represents the potential total demand after spatial distance friction decay.

The second step of the algorithm shifts the measurement perspective, taking the centroid of the predicted grid as demand point i, and back-collects all kindergarten supply point resources reachable within the time threshold t_0. The effective supply-demand ratio R_j obtained in the first step is weighted and accumulated to obtain the accessibility score A_i for each spatial unit to preschool education resources:

$$A_i = \sum_{j\in\Omega_i} R_j f(t_{ij}) \tag{6}$$

In the formula, t_{ij} represents the road network travel time from demand point i to each supply point j, $\Omega_i = \{j \mid t_{ij} \leq t_0\}$; the physical meaning of A_i is the per capita available degree equivalent within the grid. To avoid abrupt drops in resource accessibility caused by a single threshold edge effect, the model uses a continuously smooth Gaussian equation as the distance decay function $f(t)$:

$$f(t) = \begin{cases} \dfrac{\exp\left[-\frac{1}{2}\left(\frac{t}{t_0}\right)^2\right] - \exp\left(-\frac{1}{2}\right)}{1-\exp\left(-\frac{1}{2}\right)}, & t \leq t_0, \\ 0, & t > t_0. \end{cases} \tag{7}$$

t represents the travel time parameter between any two points. Through the above bidirectional spatial iterative calculation, the system transforms the scattered facility capacity and gridded population scale into a continuous two-dimensional spatial accessibility index matrix.

As shown in Figure 3, this matrix eliminates the errors of traditional Euclidean distance calculations, directly mapping the results to the map rendering layer to generate an intuitive heatmap of preschool accessibility at the street scale. In the figure, blue triangles represent existing kindergarten facilities (S_j); based on acces sibility scores (A_i), the spatial grid exhibits a clear tiered differentiation: green areas represent “resource-sufficient areas,” with an average available place score above 1.4, highly concentrated around the facilities; yellow bands represent transitional zones barely reaching the overall guarantee baseline (0.85); while dark red areas precisely expose the system’s “mismatch blind spots,” where accessibility scores plummet to an extremely scarce state of 0.2 due to distance from facilities and the superimposed

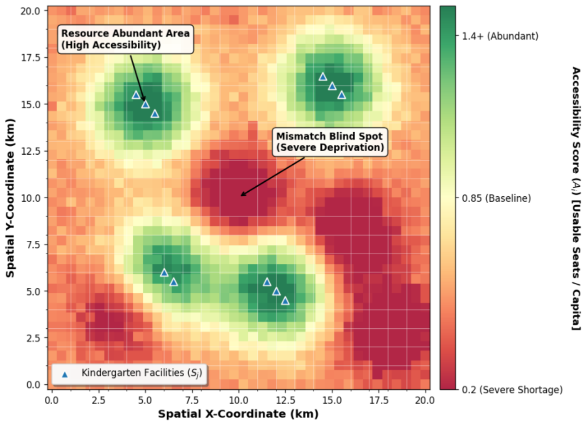

FIGURE 3. Heat map of accessibility of preschool education space at the street scale

complex spatial impedance. This quantitative spatial rendering result not only reveals the severe supply-demand mismatch in local areas but also provides high-confidence prior data and decision constraints for the subsequent multi-objective particle swarm optimization (MOPSO) algorithm to target and fill the gaps.

To achieve seamless integration of the supply-demand assessment module within the system, this i2SFCA algorithm is integrated into the system's backend spatial computing engine. During algorithm execution, the system automatically retrieves dynamic demand prediction results from the module output and existing road network and facility vector data from the PostGIS spatial database. By setting automated scheduling tasks, the system transforms high-concurrency spatial topology traversal into batch processing jobs. The calculated gridded spatial accessibility scores and mismatch blind spot coordinates are pushed in real-time to the GeoServer middleware and dynamically published as standard map tile services (WMS/WFS) for the frontend WebGL rendering engine to call and generate visual heatmap layers.

SITE SELECTION AND ALLOCATION OPTIMIZATION BASED ON MULTI-OBJECTIVE PARTICLE SWARM OPTIMIZATION

After obtaining the micro-demand dynamic evolution prediction based on the spatiotemporal graph convolutional network (STGCN) and the supply-demand mismatch prior input based on the improved two-step search algorithm (i2SFCA), the system uses the multi-objective particle swarm optimization algorithm to solve the site selection decision and capacity allocation problem of newly built and expanded kindergartens [32]. The core of the algorithm is to abstract the complex spatial resource allocation into a multi-dimensional continuous and discrete hybrid optimization constraint space, and to focus on introducing the accessibility mismatch blind zone locked in the pre-evaluation stage, the candidate points of available construction land, and the local fiscal budget as core computational constraint elements. To bridge the gap between diagnosis and construction feasibility, the MOPSO module integrates a Land-Use Suitability Layer (LUSL) as a hard spatial constraint. Candidate sites are filtered not only by their accessibility scores but also by their urban planning attributes, such as slope gradient, existing industrial zoning, and proximity to ecological redlines. This ensures that "dark red" mismatch areas are only prioritized for new facilities if the underlying land-use logic permits construction, transforming the system from a theoretical diagnostic tool into a practical architectural deployment guide. The set of available candidate sites for construction land within the city is defined as $U = \{u_1, u_2, \ldots, u_m, \ldots, u_M\}$, corresponding to two decision variable matrices: a Boolean variable $\delta_m \in \{0, 1\}$ representing whether a park should be built at candidate site u_m; and a continuous variable κ_m representing the newly allocated school place capacity at that site. Addressing the low-score spatial and physical resource constraints identified by the system, the model constructs a dual objective function that balances spatial equity and travel efficiency, and its weighting parameters are rigorously defined mathematically. The spatial equity objective, F_{equity}, aims to maximize the service coverage baseline in extremely resource-scarce areas, defined as maximizing the minimum value of the newly added accessibility score across all grids:

$$F_{\text{equity}} = \max \left(\min_{i \in \Omega_{\text{blind}}} \Delta A_i \right) \tag{8}$$

In the formula, Ω_{blind} represents the set of low-accessibility mismatch blind spots defined in the evaluation phase; ΔA_i is the marginal degree increment score obtained by grid i after expansion configuration via candidate points.

The resource efficiency objective $F_{\text{efficiency}}$ focuses on minimizing the total fiscal construction cost and the physical impedance of inter-regional travel, and its mathematical expression is:

$$F_{\text{efficiency}} = \min\Bigg(\sum_{m=1}^{M} (\eta_{\text{fix}}\delta_m + \eta_{\text{var}}\kappa_m) + \lambda \sum_{i=1}^{N} \sum_{m=1}^{M} \delta_m \cdot d_{im} \Bigg) \tag{9}$$

η_{fix} is the baseline cost of fixed facility construction; η_{var} is the marginal expansion cost per degree; d_{im} is the spatial distance from the centroid of the demand grid to the candidate point u_m; and λ is the conversion penalty coefficient for balancing fiscal expenditure and resident travel network costs. Simultaneously, the model is strictly constrained by a hard constraint of the total budget limit, meaning the total global construction cost cannot exceed the preset fiscal expenditure ceiling.

To solve the aforementioned multidimensional optimization problem with game-theoretic conflict, the MOPSO algorithm encodes each global configuration scheme as an independent particle q in a high-dimensional search space. Its position vector $\Theta_q = [\delta_1, \kappa_1, \ldots, \delta_M, \kappa_M]$ is used for adaptive position updates during iteration, based on the individual historical optimal solution L_q and the global non-dominated solution G_{best} in the external archive. The update equation for the particle's flight velocity is set as follows:

$$V_q^{(\text{iter}+1)} = \omega V_q^{(\text{iter})} + \phi_1\xi_1\Big(\vec{L}_q - \Theta_q^{(\text{iter})}\Big) + \phi_2\xi_2\Big(G_{\text{best}} - \Theta_q^{(\text{iter})}\Big) \tag{10}$$

In the formula, ω is a dynamically decreasing inertia weight used to regulate the global search breadth in the early and mid-stages and the local mining depth in the later stage; ϕ_1 and ϕ_2 correspond to the particle's cognitive and social learning factors, respectively; ξ_1 and ξ_2 are random perturbation terms following a uniform distribution in [0,1]; iter is the current iteration batch. After the algorithm reaches the convergence threshold, the system will output the Pareto optimal solution set, and finally automatically generate the optimal configuration scheme containing specific spatial location and precise construction scale. Given the high computational consumption characteristics of multidimensional continuous and discrete hybrid optimization, the MOPSO location selection and allocation module adopts an asynchronous task queue deployment mode in the system architecture. When the terminal decision-maker triggers the "optimization simulation" command on the front-end interface and sets hard constraints such as the fiscal budget, the background computing cluster immediately starts a concurrent particle swarm for iterative solution. After calculation, the system automatically extracts the non-dominated solution set from the Pareto front and returns the optimized site selection coordinates, expansion capacity allocation, and global evaluation indicators to the front-end interaction layer via JSON (JavaScript Object Notation) data stream, supporting dynamic comparison of multiple schemes and redrawing of the decision-making sandbox. To bridge these theoretical mathematical optimizations with practical administrative workflows, the system transitions from backend computation to a high-fidelity visual interface. This transition ensures that the complex multi-objective solutions derived from the MOPSO algorithm are translated into actionable insights for urban planners.

SYSTEM PROTOTYPE IMPLEMENTATION AND VISUAL INTERACTION DESIGN

Building upon the core algorithms, the system efficiently tracks the changing population trends at a micro-level, this research, based on a front-end/back-end separation architecture, microservice encapsulated the

previously verified multi-dimensional algorithm modules such as STGCN, i2SFCA, and MOPSO. Furthermore, leveraging the WebGL rendering engine, a prototype interface for the "Big Data Decision Support System for Balanced Allocation of Preschool Education Resources" was designed and implemented. The system achieves a paradigm shift from "experience-based governance" to "digital intelligent governance," and its core inter active flow and decision support functions are illustrated in Figure 4.

First, the system's main interface constructs a comprehensive data situational awareness dashboard. Through the side navigation bar, decision-makers can access the underlying spatiotemporal database in real time, intu itively overlaying multimodal data layers onto the central WebGIS map engine. For example, after calling the STGCN prediction module, the map area will automatically render the evolution trajectory of the population center of gravity of school-age children and high-density overflow areas over the next few years in the form of a dynamic heatmap. Secondly, the system contains an interactive mismatch blind spot diagnosis pop-up window. Within the supply and demand assessment module, decision-makers will be able to customize the extreme threshold travel time (for example 15 minutes). The i2SFCA algorithm will run in real time in the background of the system and will graphically highlight a mismatch blind spot on the map (for example, with a bright red block). The data panel on the right will simultaneously pop up with the specific capacity gap, the current per capita score of Accessibility, and the impedance characteristics of the road network in that area, therefore turning the abstract mathematical calculation into an intuitive "diagnostic report". Finally, the system integrates extensively the multi-objective site selection/allocation sand table simulation functionalities. Administrative decision-makers can enter their hard constraints (for example, their current "fiscal budget ceiling") into the parameter control panel located at the top of the interface. After initialising MOPSO, the system will provide a dynamic plot of the Pareto curve at the bottom of the panel, and will automatically anchor the optimal solution to the map of new kindergartens and generate recommended physical coordinates for new kindergartens/expansion capacity (marked with indicative icons

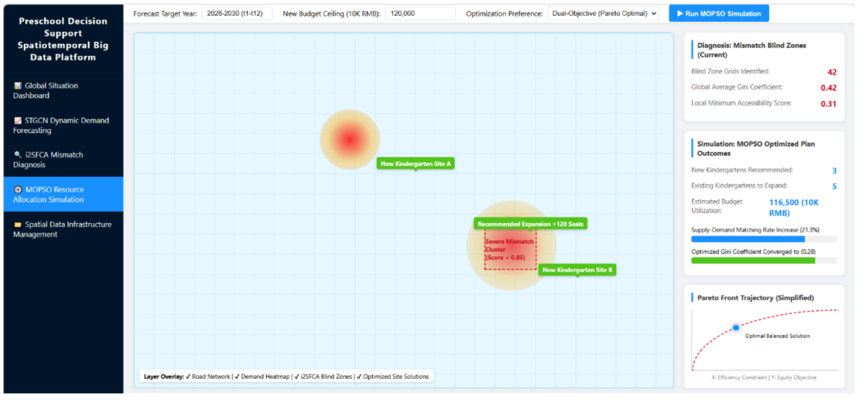

FIGURE 4. System the matrix D isplay and Data Visualization

like stars). The comprehensive control panel to the right will calculate and provide real-time comparative values of the global Gini coefficient and Matching Degree Improvement before and after the optimisation process, therefore generating a high degree of confidence in education resource allocation (macro-level) through system-level evidence.

III. EXPERIMENTAL DESIGN AND DATASET CONSTRUCTION

OVERVIEW OF EXPERIMENTAL AREA AND CONSTRUCTION OF MULTI-SOURCE DATASETS

To examine the effectiveness of the engineering of a system of system, a typical high-density urban area has been selected as a study area for the experimental program. The study area is about 150 km2 in size and has an old residential neighbourhood, new commercial housing, and a commercial area. This large variance of population density and significant fluctuations in local demand meet all of the real world needs of the system because they demonstrate the ability of the system to adapt to dynamic demand environments. At the level of spatial reference construction, the WGS-84 (World Geodetic System 1984) coordinate reference system was used to define the spatial reference frame. The study area has been subdivided into discrete standard 500 m $\times$ 500 m spatial units using vector gridding technology, thereby producing 600 applicable computational grids to be used as the lowest level of granularity for subsequent predictions of micro-demand and resource allocation.

The system used an operator-controlled approach that allowed it to extract multi-dimensional heterogeneous data from 60 months of consecutive data (January 2019 through December 2023) for model training and testing. The collection of the multi-source heterogeneous data required rigorous data cleansing and alignment processes (i.e., fusion/preprocessing) in accordance with the standards outlined in the method section. For any sporadic missing data elements in the population time series, a linear interpolation method was applied to complete these data elements; while an outlier removal (moving average) filter (i.e., three standard deviations above mean) was applied to any high population spikes in order to smooth the data. Spatially, the use of a polygon overlay analysis tool created a grid-based projected area that uniformly coordinates all independent kindergarten POIs and vector road segments on one grid system. Topologically, a graph traversal algorithm produced corrections to the original street network data by eliminating dangling nodes and isolated subgraphs; thereby providing the generated spatial adjacency matrix (used in the reachability calculations) with a mathematically determining procedure for achieving strictly positive definite connectivity. As per the table below, the specific characteristics, depths, and relationships of the four (4) core fields (micropopulation grid, POIs for education facilities, transportation networks, and urban planning land use) are thoroughly represented across the multi-source spatiotemporal dataset's four (4) core dimensions (micropopulation grid, POIs for education facilities, transportation networks, and land use).

TABLE 1. Description of Features and Core Fields of the Multi-Source Spatiotemporal Dataset

Dataset Type	Data Source	Data Scale / Span	Core Field Descriptions
Micro-level Population Grid	Public Security Household Registration and Mobile Signaling records	60 months / 36,000	Grid_ID, Time_Stamp, Data Pop_Count (ages 0–6)
Kindergarten POI Data	Education Administrative Approval System	214 facilities	KG_ID (Kindergarten Identifier), Longitude/Latitude, Capacity (seats)
Traffic Road Network Vector Data	Planning and Natural Resources Bureau	15,000+ road segments	Node_Start/Node_End, Length (m), Speed_Limit (km/h)
Urban Planning Land Use Data	Regulatory Planning Database	32 candidate plots	Land_ID, Center_Coord, Buildable_Area (m2)

PREDICTION AND EVALUATION MODEL PARAMETER SETTINGS AND BASELINE MODEL SETTING

A computer with two NVIDIA RTX 3090 GPUs (Graphics Processing Units) and 128 GB of memory serves as the hardware environment for the experimental engineering platform. PyTorch 2.0 serves as the underlying computing framework. The front-end interface for system prototype development and visualization uses the Vue.js progressive framework and utilizes Mapbox GL and ECharts to render WebGIS spatial data, enabling high concurrency, dynamic heatmap generation, vector road network loading and millisecond level response on multiple decision indicator panels.

To rigorously verify the performance advantages of the Spatiotemporal Graph Convolutional Network (STGCN) in dynamic correlation feature extraction, the experiment constructed a stepped baseline model matrix from basic statistics to deep time series for quantitative comparison:

HA (Historical Average): As a naive benchmark, it directly calculates the arithmetic mean of the population of each grid within the same historical time window, used to assess the normal periodicity of the data.

ARIMA (Autoregressive Integral Moving Average Model) [33]: Linear fitting is performed independently on the population time series of each basic spatial grid. The grid-level parameter combination search is performed using the Akaike Information Criterion (AIC) to finally determine the optimal discrete distribution matrix of the difference order, number of autoregressive terms and number of moving average terms (p, d, q).

LSTM (Long Short-Term Memory Network) [34]: As a representative of deep learning time series models, it removes the influence of spatial topology dimension and focuses on the non-linear feature extraction of singlegrid historical time series. Its hidden layer dimension is set to 64, and the time step is consistent with STGCN. The aforementioned control group configuration aims to systematically and quantitatively evaluate the improvement in error convergence and generalization ability of the spatial topology algorithm compared to traditional pure time series analysis.

For the instantiation of the micro-demand dynamic prediction module, the hyperparameters of the STGCN network structure are strictly aligned with the data base granularity. The input time series sliding window length is truncated to 12 time steps to accurately extract the annual periodic oscillation features of population migration and fertility patterns. The spatial dimension stacking is set to a 3-layer graph convolutional network to ensure that the model effectively captures the spillover effect of at most three-order neighborhoods of grid nodes. During the model training phase, the Adam optimizer is specified to perform gradient descent iterations, with an initial learning rate fixed at 0.001, while a constant-level weight decay coefficient is applied to suppress the overfitting risk caused by high-dimensional features. The loss evaluation function uniformly adopts mean squared error, the total number of iterations is set to 200, and an early stopping trigger mechanism based on the validation set loss not decreasing for 15 consecutive rounds is forcibly embedded. In the resource scheduling and location optimization phase, to highlight the superiority of the MOPSO dual-objective collaborative mechanism, the experiment constructed a baseline matrix for resource allocation tasks on an evaluation base with a maximum travel time threshold t_0 of 15 minutes. Besides retaining the current static configuration as the original mismatch baseline, the model decomposed the objective function into control variables, instantiating two control groups: a single-objective efficiency-first configuration (activating only the $F_{\text{efficiency}}$ objective function) and a single-objective fairness-first configuration (activating only the F_{equity} objective function).

Both the single-objective control groups and the proposed MOPSO dual-objective algorithm were concur rently extrapolated under the same hard constraint of total budget limit. The MOPSO algorithm's population size was initialized to 100 concurrent particles, and the maximum number of iterations was strongly constrained to 500. To avoid the algorithm getting trapped in local optima, the inertia weight ω adopted an adaptive decreasing strategy, linearly decaying from 0.9 to 0.4 with the iteration process; the cognitive learning factor and the social learning factor ($\phi 1$, $\phi 2$) were both statically assigned a value of 2.0. The specific engineering hyperparameter configurations and strategies for each core algorithm have been systematically organized and summarized in Table 2.

TABLE 2. Hyperparameter Configuration Table for Prediction and Optimization Core Algorithms

Algorithm Module	Core Parameter	Setting / Strategy	Engineering Significance
HA (Baseline)	Time Window	12 months	Extracts normal arithmetic mean of historical periods
ARIMA (Baseline)	p, d, q	AIC-based Dynamic Optimization	Controls autoregressive, differencing, and moving average terms
LSTM (Baseline)	Hidden Size	64	Controls representation dimension of non-linear temporal features
STGCN	Time Window Length	12 steps	Determines historical time span for model reference
STGCN	Graph Convolution Layers	3	Controls aggregation depth of spatial topological neighborhood
STGCN	Learning Rate	0.001	Regulates step size for loss function gradient descent
STGCN	Optimizer	Adam	Implements adaptive learning rate adjustment and parameter iteration
i2SFCA	Time Threshold (t_0)	15 min	Defines physical boundary for effective preschool resource coverage
MOPSO	Swarm Size	100	Determines concurrent solution vectors in multi-objective search space
MOPSO	Max Iterations	500	Controls termination condition and computational lifecycle
MOPSO	Inertia Weight (ω)	$0.9 \rightarrow 0.4$ (Linear Decay)	Balances global exploration and local exploitation capabilities
MOPSO	Learning Factors (ϕ_1, ϕ_2)	2.0, 2.0	Adjusts acceleration weights toward individual and global optima

IV. RESULTS ANALYSIS

VALIDATION OF STGCN MICRO-DEMAND DYNAMIC PREDICTION ACCURACY

In order to calculate the advantage in performance of the use of spatiotemporal graph convolutional networks (STGCN) for predicting micro-scale school-age populations, an experiment was performed where the indepen dent test dataset consisted of gridded observation data from 12 months (consecutively) in 2023. Additionally, the predicted values for school age children obtained through the STGCN model (those considered under the STGCN model) were compared with the outputs of baseline models such as HA, ARIMA, and LSTM. All predictions from the STGCN model were generated using only 500m $\times$ 500m as the computational unit. Therefore, MAPE and RMSE (mean absolute percent error and root mean square error, respectively) were calculated for each grid as part of the validation process.

As evidenced by both empirical data and error metric comparisons for various prediction algorithms, Figure 5 shows a clear stepwise increase in prediction accuracy as the model can better extract spatio-temporal features. The HA model (naive benchmark) had the highest error rate with a global MAPE and RMSE of 25.5% and 35.2% respectively. The traditional ARIMA model used an autoregressive mechanism based on a single grid time-series, resulting in a global MAPE of 21.3%. The deep-learning time-series baseline LSTM model was more capable of non-linear fitting, resulting in a global MAPE of 12.1%. Finally, with respect to the STGCN's architecture, it produced significantly lower global MAPE at 6.8% and produced grid-level RMSE that was nearly converged at 12.4% (i.e., an absolute discrete bias dimension). All error metrics achieved their most optimal level in this case. This relative performance establishes mathematically that joint spatiotemporal solutions have a key effect on controlling the variance associated with predictions and increasing the global accuracy of predictions.

When dealing with dramatic fluctuations in micro-population, pure time-series models (such as ARIMA and LSTM) rely solely on the historical sequence of a single spatial unit, making it difficult to effectively quantify cross-regional population agglomeration caused by the delivery of new residential areas or urban renewal. Therefore, significant biases are still evident in the global error statistics shown in Figure 5. In contrast, the STGCN model, relying on graph convolution operators to synchronously aggregate multi-level topological neighborhood information, successfully absorbs and internalizes the population spillover characteristics generated by adjacent high-density grids towards the target grid. The significant drop in the STGCN error histogram in Figure 5 is a direct reflection of this agile spatial feature capture capability in terms of global accuracy. This process accurately maps the complex demand evolution trends in real physical space, completely overcoming the spatial planning lag caused by the high dependence of traditional resource allocation models on static household registration bases. Through the engineering implementation of this algorithm module, the system successfully outputs highly confident micro-demand indicators when processing dramatically fluctuating population base data.

PRECISE IDENTIFICATION OF EXISTING SPATIAL MISMATCH BLIND SPOTS

The predicted sequence of school-age population in the micro-grids output by the STGCN model is matrix-aligned with the approved capacity of existing kindergartens in the underlying spatial database. Based on the improved two-step search algorithm (i2SFCA), the system sets a 15-minute actual road network walking distance as the extreme time threshold, calculating the per capita available school place equivalent score under dynamic demand for each grid. The calculation of the actual "mismatch blind spots" from vast quantities of spatial data has been accomplished through the use of an algorithm that has used a very strict composite judging criterion – A location has been judged to be a

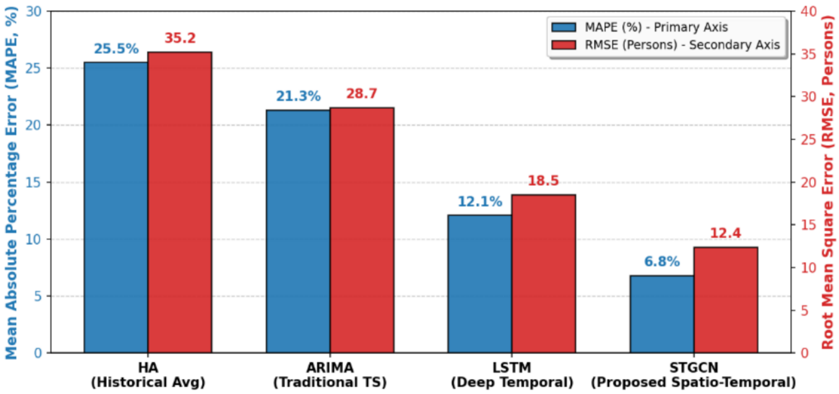

FIGURE 5. Bar chart comparing error metrics of different prediction algorithms

location with a significant mismatch with respect to its surrounding area if the accessibility score (A_i) for that location is less than the average guarantee line for the total area under consideration during three consecutive projected periods and the total local absolute increase in demand (ΔD_i) at that location and in its associated topological area exceeds a total number of 50 persons. The mathematical expression of this composite judgment condition is expressed as follows:

$$A_i < 0.85 \quad \text{and} \quad \Delta D_i > 50 \tag{11}$$

After traversing and scanning 600 spatial computing grids across the entire domain and performing polygon overlay analysis, the system accurately identified 42 hidden supply-demand mismatch blind spots that were missed by traditional empirical models.

As shown in Figure 6, the system seamlessly pushes the coordinates of these 42 blind spots and the quantified gap sequence to the GIS rendering engine, transforming them into a visualization layer with realistic physical boundaries and road network topology. The spatial clustering calculation results are visually confirmed in two typical magnified views in Figure 6:

In the western old city view (Figure 6a), the system detected 15 local scarcity points exhibiting “resource island” characteristics. Although the overall working-age population in this area is shrinking, the complex network of one-way streets and extremely high physical traffic impedance (as shown by the irregular gray lines in the background layer) severely restrict the grid’s ability to obtain effective resources from surrounding existing facilities (S_j).

In the newly developed industrial park window in the east (Figure 6b), due to the spillover effect of the rapid influx of residents, 27 grids exhibit a phased shortage of school places. Under traditional static household registration planning, this area was once assessed as “resource-compliant.” However, under dynamic demand evolution, even with a well-organized gridded road network, the rapid influx of population has created a highly concentrated cluster of severe accessibility gaps, with the actual per capita accessibility score plummeting to 0.31.

The aforementioned cross-validated blind spot identification results concretize the abstract supply-demand mismatch into the real urban mechanism, directly providing a high-confidence prior constraint boundary for the subsequent targeted optimization of the multi-objective particle swarm optimization algorithm (MOPSO).

EVALUATION OF THE IMPLEMENTATION EFFECT OF THE SYSTEM’S AUTOMATIC CONFIGURATION SCHEME

To verify the global allocation efficiency of the output scheme of the Multi-Objective Particle Swarm Optimization (MOPSO) algorithm, the experiment injected the system’s recommended non-dominated solution set (site coordinates and expansion capacity) as independent incremental parameters into the underlying spatial database, and re-scheduled the improved two-step move-search (i2SFCA) module to perform global computational power extrapolation. To highlight the superiority of the multi-objective collaborative mechanism, traditional static configuration, single-objective efficiency-first configuration, and single-objective fairness-first configuration were introduced as control groups during the evaluation phase. Under the hard constraint of strictly anchoring the upper limit of the local government’s new budget, the system performed a secondary calculation and matrix comparison of the per capita accessibility score (Ai) of 600 standard grids across the entire domain under dynamic demand. This calculation process completely eliminated subjective experience intervention and, relying on a unified spatiotemporal benchmark, achieved objective data restoration of the implementation efficiency of complex allocation schemes.

Based on the aforementioned high-dimensional comparison matrix, the discrete and central tendency statis tics of the global score sequence were extracted for verification, as shown in the box plot of the global spatial accessibility score in Figure 7. Quantitative characteristics show that under the current static configuration, the lower quartile (Q_1) of the global grid score is only 0.45, exhibiting a severely right-skewed long-tailed polarized distribution. When adopting a single-objective efficiency-first configuration, although the

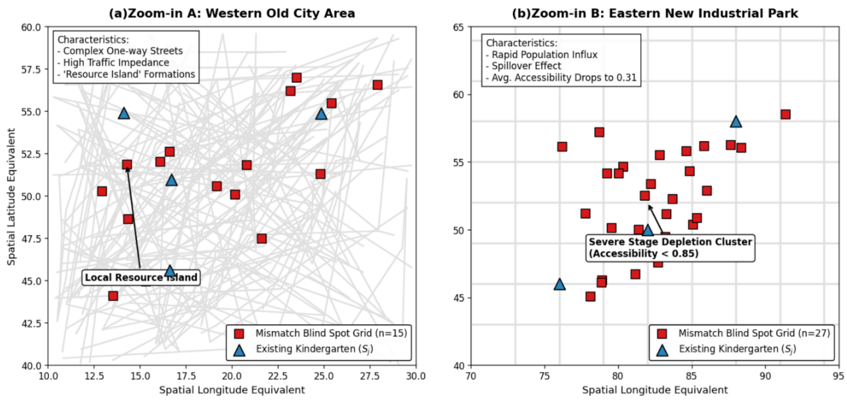

FIGURE 6. GIS-enlarged map of resource supply and demand mismatch in key gap areas: (a) Western old city; (b) Eastern new industrial park

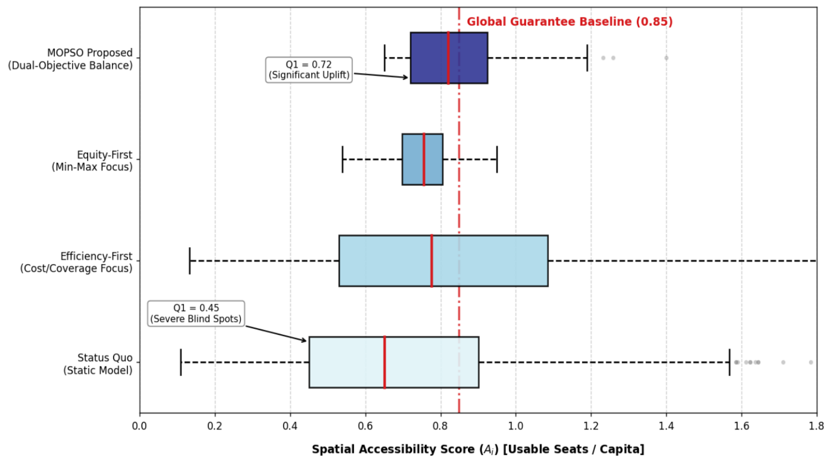

FIGURE 7. the matrix D istribution of global spatial reachability scores across optimization strategies

global mean improves, resources are heavily skewed towards low-cost, high-density areas, and Q_1 remains low at 0.53, failing to effectively fill the original "mismatch blind spot." If a single-objective fairness-first configuration is adopted, although spatial polarization is forcibly deconstructed, due to budget constraints, a large amount of funds is spent on remote, high-impedance areas, resulting in a severely suppressed global mean; the overall box plot even fails to reach the guaranteed threshold of 0.85.

In contrast, after implementing the MOPSO dual-objective collaborative optimization configuration generated by this system, the Q1 score of the global grid accessibility score jumped significantly to 0.72, and the lower limit of the score (excluding outliers) approached the global average guarantee line, confirming that the system implemented high-precision targeted filling of the original "mismatch blind spots." Simultaneously, the overall distribution of spatial accessibility transformed into a highly convergent, concentrated distribution. The overall increase in the underlying benchmark data and the sharp decrease in global variance confirm that this decision support system, in revitalizing existing educational resources and precisely allocating incremental resources, effectively mitigated inter-regional allocation disparities with minimal public financial expenditure, completely breaking the limitations of single-objective optimization and the resource allocation barriers caused by lagging spatial planning.

ANALYSIS OF THE EVOLUTION OF DEGREE SUPPLY AND DEMAND MATCHING AND SPATIAL FAIRNESS

To deeply quantify the evolution of spatial equity in resource allocation under dynamic demand shocks, a Lorenz curve was constructed based on the grid-level accessibility score matrix output by an improved two-step moving search method, representing the cumulative proportion of the school-age population and the cumulative proportion of available school places. By calculating the area enclosed by the absolute equity line and the actual allocation curve, the Gini coefficient of preschool education resources under different allocation schemes within the continuous prediction period t1 to t12 was accurately obtained.

As noted in Figure 8 (a), there are many differences in how four strategies address population spatial spillover when comparing their performance using comparative time series analysis. Under the current static arrangement, the basic Gini coefficient for the experimental subject area was originally 0.42 at the beginning of the series, but it has continued to deteriorate with each subsequent time period due to a failure to adjust to the increase in demand for school places in that subject area. This is evidenced by the Gini coefficient consistently exceeding the warning threshold (0.40) and eventually reaching 0.49, suggesting further increases in both spatial polarization and localized monopolies in the distribution of facilities. Excessive concentration of funding in high-density population areas for efficiency-first allocation, as described above, will result in an overall decrease in the value of the Gini coefficient, however, its value will continue to be high and fluctuate between 0.36 and 0.38 for the duration of the forecast period approaching the warning threshold, resulting in continued significant levels of marginal deprivation. Conversely, while the single-objective fairness-first allocation sacrifices efficiency to flatten the Gini coefficient to a low level of 0.21 to 0.23, it lacks robustness against fluctuations and fiscal sustainability in a real and complex dynamic environment. After the MOPSO dual-objective collaborative optimization configuration proposed in this system is injected, the global Gini coefficient steadily decreases from the initial 0.30 and converges stably to the ideal range of 0.28, which is within the "Optimal Balance Zone" (0.25-0.35) shown in Figure 8a. This dynamic evolution process confirms from a macroscopic mathematical perspective that the algorithm's built-in dynamic balance mechanism between fairness and efficiency plays a dominant role, successfully driving incremental resources to tilt towards the dynamically evolving edge grids, and substantially dismantling the original structural imbalance barrier.

In the verification of the micro-level supply and demand matching degree (as shown in Figure 8b), the algorithm extracts the dynamic demand evolution tensor $D^{(t)}$ output by the STGCN module under continuous time window, and aligns it with the ultimate carrying capacity matrix S of each spatial supply point (including existing retention and optimized addition) across cycles. Using the set ultimate impedance threshold t_0 as a rigid constraint boundary, the absolute ratio of the total available supply within the effective radiation range to the actual predicted total population is calculated grid by grid. Its mathematical evaluation logic is expressed as follows:

$$M_{\text{ratio}} = \frac{\sum_j S_j}{\sum_i D_i(t)} \tag{12}$$

Figure 8b's data tracking and differential calculation intuitively show that, under the impact of dynamic population influx, the matching degree of the current static configuration continuously declined from 0.72 to a local shortage state of 0.66; however, through the closed-loop optimization and allocation of the decision-making system, the MOPSO configuration scheme not only rapidly broke through the overall guarantee baseline of 0.85, but also stabilized at 0.89. Comparing the two, the overall supply and demand matching degree across the entire region achieved a quantitative leap of 23% compared to the current situation. This improvement in the underlying logic of the indicator completely severed the mismatch chain caused by the "static indicator allocation" in traditional empirical models, which resulted in both local vacant school places and overcrowding in schools. Under the spatiotemporal coupling of physical mapping, it rigorously verified the engineering effectiveness of the system's agile response to dynamic population fluctuations.

MULTI-OBJECTIVE PARETO FRONT ANALYSIS OF SITE SELECTION AND CAPACITY ALLOCATION

After running 500 full-load iterations, the MOPSO algorithm found a total of 500 global nondominated solutions and plotted them on a two-dimensional decision space defined by the spatial equity objective F_{equity} and the resource efficiency objective $F_{\text{efficiency}}$ serving as the two coordinate system axes. The continuation of non-dominated solution vector has produced a characteristic non-linear convex-shaped Pareto Front (shown in red on the performance plots) and is significantly better than large scale historical nondominated (gray) results within the search area. On this Pareto Front line, for each increment in performance in a single performance dimension, there is a corresponding decrement (or loss) of performance in another performance dimension and demonstrates from a mathematical perspective that there is a strong conflict between the indicators of equity and efficiency for selecting sites for preschool education and distributing available capacity. Tracing the boundary extreme points on the right side of the curve reveals that when the system algorithm attempts to pursue the ultimate improvement in spatial accessibility, the required marginal fiscal construction cost per degree and the total time spent on cross-regional commuting networks both increase exponentially.

To overcome the decision-making dilemma caused by multidimensional non-dominated solution sets, the system incorporates a dynamic inflection point discrimination mechanism based on the Utopia Point (marked by a yellow star in Figure 9), performing secondary geometric feature extraction on the Pareto front. The algorithm finds the globally optimal recommended solution, which is marked by blue circles in Figure 9, by evaluating and identifying point on the curve which has the closest distance to the theoretical absolute optimum rating of both goals. To do this, the normalized Euclidean distance from each feasible solution vector along the curve is computed to the theoretical absolute optimum for both goals (the theoretical intersection of minimum cost and maximum fairness).

The optimal solution contains the decoded parameters which show that, while strictly meeting the hard constraint of increasing local tax allowable amount, the proposed solution not only brings the core access score of very poorly matched grids very close to the guaranteed red line (0.85), it also eliminates redundancy in degrees of

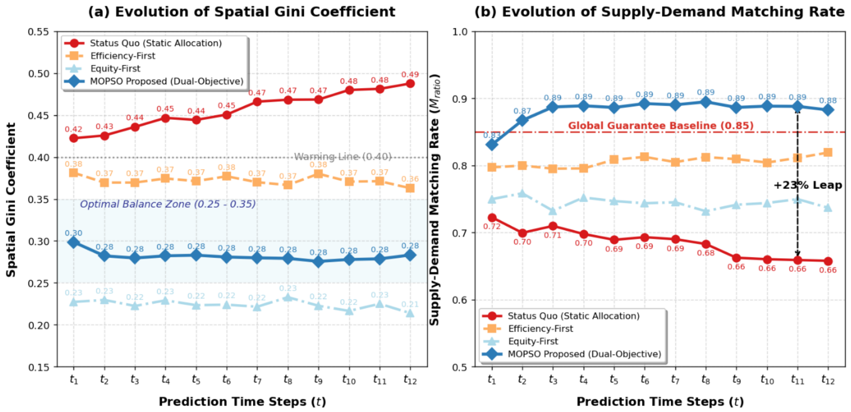

FIGURE 8. Line chart comparing spatial Gini coefficients of different optimization strategies under dynamic demand evolution: (a) Evolution of spatial Gini coefficient; (b) Evolution of supply-demand matching rate

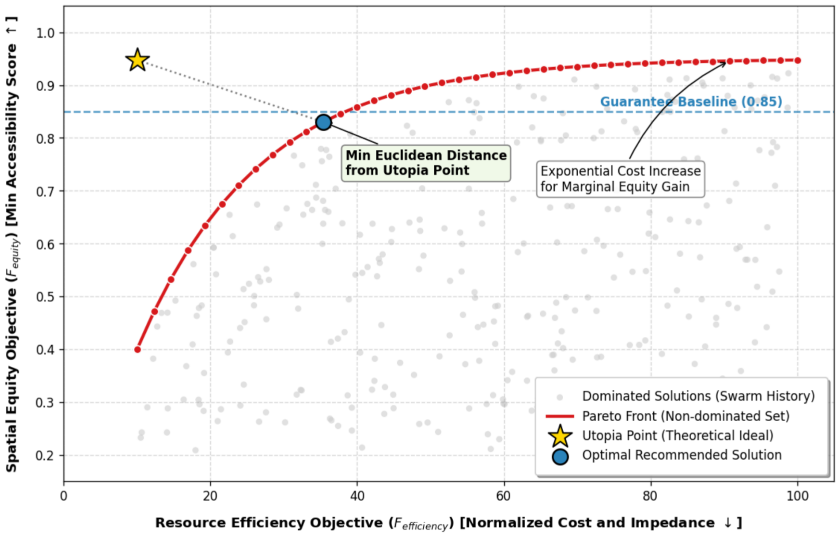

FIGURE 9. Pareto frontier distribution of MOPSO site selection and capacity optimization

access and inefficient use of funds associated with blindly tilting toward remote, low population density areas through smarter capacity allocation. This entire multi-objective analysis process has completely transformed the path of the heuristic algorithm's quest for an optimal solution in a clear and transparent way, by transforming the combined constraints of interdependence into quantitative allocation evidence which supports macroeconomic policy decisions.

V. CONCLUSION

To address the supply-demand mismatch and lagging spatial planning caused by drastic population fluctuations within manufacturing clusters, this paper implements a spatiotemporal big data decision support system for the balanced allocation of preschool education resources. In terms of methodology and system engineering implementation, this research establishes a complete closed loop from underlying algorithm derivation to front-end interactive decision-making. First, the system constructs a multi-source data fusion foundation and introduces a spatiotemporal graph convolutional network to accurately characterize the evolution of demand and spatial spillover effects of preschool-age children at the micro-block scale. Second, by combining an improved two-step moving search method and a multi-objective particle swarm optimization algorithm, a facility site selection and capacity dynamic allocation model that balances spatial equity and travel efficiency is constructed. Based on this, an interactive front-end prototype is developed using the WebGL rendering engine, realizing the visualization integration of full-domain situational awareness, dynamic diagnosis of mismatch blind spots, and a multi-dimensional auxiliary decision-making sandbox. Experimental and application results on real-world urban datasets demonstrate that, at the algorithmic level, the STGCN model reduces the average absolute percentage error of micro-demand prediction to 6.8%. Compared to traditional static allocation models, the system's output collaborative optimization scheme effectively lowers the Gini coefficient for global spatial accessibility, significantly improves the spatial fairness of resource allocation, and increases the degree supply-demand matching rate by 23%. At the application level, the system's automatically output spatiotemporal heatmap and Pareto front allocation index effectively transform abstract macroscopic theoretical optimal solutions into visualized inference evidence with real-world feasibility. In conclusion, this system overcomes the limitations of traditional experience-based planning by synchronizing public resource distribution with the demographic shifts inherent in manufacturing agglomerations, providing government departments with a high-confidence methodology for agile and precise 'digital intelligent governance'.

AUTHOR CONTRIBUTIONS

Conceptualization – Jiaxin Chen, Yang Yang and Wensheng Yan; methodology – Jiaxin Chen, Yang Yang and Wensheng Yan; investigation – Jiaxin Chen, Yang Yang and Wensheng Yan; writing-original draft preparation – Jiaxin Chen, Yang Yang and Wensheng Yan. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] J. Sun, H. Wu, and S. Shi, "A research of the evaluation of preschool education resource allocation level and spatiotemporal differences: Based on repeated indicators method and Theil index," Heliyon, vol. 9, no. 6, pp. 1-15, 2023, doi: 10.1016/ j.heliyon.2023.e16362.

[2] Y. Wang, J. Calero, C. Gao, et al., "Temporal and spatial dynamics of preschool education resource allocation in China from 2011 to 2019: Based on statistical and spatial analysis methods," Sage Open, vol. 14, no. 4, pp. 1-14, 2024, doi: 10.1177/21582440241289689.

[3] L. Rong, F. Deng, Z. Zhong, et al., "Impact Analysis of Chinese Government Leadership in the Construction of a Public Service System of Preschool Education," Education and Urban Society, vol. 56, no. 7, pp. 881-905, 2024, doi: 10.1177/00131245241226671.

[4] H. Li, "Research on educational space design for preschool children," Journal of Theory and Practice of Social Science, vol. 3, no. 11, pp. 19-27, 2023, doi: 10.53469/jtpss.2023.03(11).04.

[5] F. Pang, Y. Li, G. Miao, et al., "Optimizing education resource allocation using grey model forecasting of school age populations," Scientific Reports, vol. 15, no. 1, pp. 1-17, 2025, doi: 10.1038/s41598-025-19182-5.

[6] M. Tang M, Z. Zeng, and Q. Lan, "How to evaluate the efficiency of rural preschool education resources and its regional differences in China," Scientific Reports, vol. 14, no. 1, pp. 1-11, 2024, doi: 10.1038/s41598-024-72892-0.

[7] P. Machebele and G. Weir-Smith, "Exploring the role of spatial mismatch on educational attainment in South African metropolitan areas: spatial data analysis," South African Geographical Journal, vol. 107, no. 2, pp. 239-261, 2025, doi: 10.1080/03736245.2024.2381019.

[8] Y. Hu, P. Sherlock, J. Huang, et al., "Unveiling spatial mismatch in childcare supply and demand: An excess commuting analysis of home-to-childcare distance in subsidized families," Journal of Transport Geography, vol. 116, no. 1, pp. 1-7, 2024, doi: 10.1016/j.jtrangeo.2024.103829.

[9] M. Tang M and Q. Lan, "Research on policy effect and regional difference of allocation efficiency of rural preschool education resources: evidence from China," Plos one, vol. 20, no. 5, pp. 1-20, 2025, doi: 10.1371/journal.pone.0301064.

[10] Z. Luo, J. Wang, and X. Bai, "Spatial the matrix D istribution and Influencing Factors of Shadow Education Enterprises in Primary and Secondary Schools: A Case Study of Zhengzhou's Main Urban Area," Frontiers, vol. 4, no. 12, pp. 1-24, 2023, doi: 10.6981/ FEM.202312_4(12).0002.

[11] Q. Huang, X. Cui, and L. Ma, "The equity of basic educational facilities from the perspective of space," Sustainability, vol. 15, no. 15, pp. 12031-12047, 2023, doi: 10.3390/su151512031.

[12] Y. Chen, Z. Lai, and C. Huang, "Optimizing spatial location and service capacity of new schools toward maximum equity in the distribution of educational resources," Journal of Urban Planning and Development, vol. 149, no. 3, pp. 1-8, 2023, doi: 10.1061/JUPDDM.UPENG-4346.

[13] C. Lu, C. Yu, and X. Liu, "Evaluating the Quality of Children's Active School Travel Spaces and the Mechanisms of School the matrix D istrict Friendliness Impact Based on Multi-Source Big Data," Land, vol. 13, no. 8, pp. 1319-1337, 2024, doi: 10.3390/ land13081319.

[14] Z. Sun, Y. Xu, J. Ning, et al., "Educational Facility Site Selection Based on Multi-Source Data and Ensemble Learning: A Case Study of Primary Schools in Tianjin," ISPRS International Journal of Geo-Information, vol. 14, no. 9, pp. 337-359, 2025, doi: 10.3390/ijgi14090337.

[15] L. Wang and T. Liu, "Optimization algorithm of preschool education resource allocation based on data envelopment analysis model," Scientific Programming, vol. 2022, no. 1, pp. 1-11, 2022, doi: 10.1155/2022/3858664.

[16] J. Huang, "Preschool education resource allocation model for index system evaluation based on non-linear random matrix," Mathematical Problems in Engineering, vol. 2022, no. 1, pp. 1-10, 2022, doi: 10.1155/2022/6271690.

[17] F. Xia, "Forecasting Teacher Demand in China's Compulsory Education: A Time Series Analysis Approach," Exploration of Humanities and Social Research, vol. 1, no. 1, pp. 1-9, 2025.

[18] X. Zhou, "A Study on the Construction of a Predictive Model for Physical Education Teachers' Professional Development Trajectories Supported by ARIMA Algorithm Driven by Physical Education Reforms," J. Combin. Math. Combin. Comput., vol. 127, no. 1, pp. 5711-5730, 2025, doi: 10.61091/jcmcc127b-315.

[19] Y. Wang, "Prediction of preschool education demand under china's three-child policy: Based on the demographic trends from 2024-2050," Early Childhood Education Journal, vol. 53, no. 5, pp. 1823-1834, 2025, doi: 10.1007/ s10643-024-01709-6.

[20] D. Yang, J. Sun, S. Xie, et al., "Spatial Accessibility Characteristics and Optimization of Multi-Stage Schools in Rural Mountainous Areas in China: A Case Study of Qixingguan the matrix D istrict," Sustainability, vol. 17, no. 9, pp. 3862-3882, 2025, doi: 10.3390/su17093862.

[21] Y. Liao and K. Furuya, "A case study on children's accessibility in urban parks in Changsha City, China: developing an improved 2SFCA method," Land, vol. 13, no. 9, pp. 1522-1540, 2024, doi: 10.3390/land13091522.

[22] M. Köse, M. Koçyiğit, C. Erdem, et al., "An evaluation of accessibility to preschool education institutions using geographic information systems," Education and Information Technologies, vol. 26, no. 4, pp. -4328, 2021, doi: 10.1007/ s10639-021-10482-7.

[23] M. Gómez J, A. Moncada C, and A. Escobar D, "Spatial equity analysis of educational service," Methodological proposal based on a transport supply model. Geo Journal of Tourism and Geosites, vol. 51, no. 1, pp. 1622-1628, 2023, doi: 10.30892/ gtg.514spl03-1158.

[24] I. Meca, G. Nicolaescu, M. Herbei, et al., "GIS SPATIAL DATABASES FOR MANAGEMENT OF URBAN GREEN AREAS," Research Journal of Agricultural Science, vol. 57, no. 3, pp. 1-8, 2025, doi: 10.59463/RJAS.2025.3.29.

[25] S. Venkatapuram S, A. Gogineni, and S. Reddy R R, "Geospatial Data Warehousing: A Case Study on PostGIS, QGIS, and Apache Solr Integration," International Journal of Research Publication and Reviews, vol. 5, no. 1, pp. 6501-6507, 2024, doi: 10.55248/gengpi.5.0424.1089.

[26] J. Duque and M. Brovelli, "System architecture for geospatial virtual data integration in web-based applications," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 48, no. 1, pp. 939-944, 2023, doi: 10.5194/isprs-archives-XLVIII-1-W2-2023-939-2023.

[27] S. Chowdhury, D. Schröder, H. Ostadabbas, et al., "Modernizing geospatial services: an investigation into modern OGC API implementation and comparative analysis with traditional standards in a web application," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 48, no. 1, pp. 11-18, 2024, doi: 10.5194/isprs-archives-XLVIII-4-W12-2024-11-2024.

[28] G. Jin, Y. Liang, Y. Fang, et al., "Spatio-temporal graph neural networks for predictive learning in urban computing: A survey," IEEE transactions

on knowledge and data engineering, vol. 36, no. 10, pp. 5388-5408, 2023, doi: 10.1109/ TKDE.2023.3333824.

[29] Z. Mo, H. Xiang, and X. Di, "Cross-and context-aware attention based spatial-temporal graph convolutional networks for human mobility prediction," ACM Transactions on Spatial Algorithms and Systems, vol. 10, no. 4, pp. 1-25, 2024, doi: 10.1145/3673227.

[30] L. Liu, H. Lyu, Y. Zhao, et al., "An improved two-step floating catchment area (2SFCA) method for measuring spatial accessibility to elderly care facilities in Xi'an, China," International Journal of Environmental Research and Public Health, vol. 19, no. 18, pp. 11465-11480, 2022, doi: 10.3390/ijerph191811465.

[31] W. Dai, S. Yuan, Y. Liu, et al., "Measuring equality in access to urban parks: A big data analysis from Chengdu," Frontiers in Public Health, vol. 10, no. 1, pp. 1-19, 2022, doi: 10.3389/fpubh.2022.1022666.

[32] Š. Hubálovský, M. Hubálovská, and I. Matoušová, "A new hybrid particle swarm optimization-teaching-learning-based optimization for solving optimization problems," Biomimetics, vol. 9, no. 1, pp. 8-38, 2023, doi: 10.3390/biomimetics9010008.

[33] L. Schaffer A, A. Dobbins T, and S.-A. Pearson, "Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions," BMC medical research methodology, vol. 21, no. 1, pp. 58-70, 2021, doi: 10.1186/s12874-021-01235-8.

[34] Y. Chen, "Evaluation of teaching effect of internet of things education platform based on long-term and short-term memory network," International Journal of Continuing Engineering Education and Life Long Learning, vol. 31, no. 1, pp. 36-52, 2021, doi: 10.1504/IJCEELL.2021.111839.