Evaluation and Management Optimization of Higher Education Resource Allocation Efficiency Driven by Big Data

Donghe Liang[1], Xiaoying Dong[1], Chan Bai[1], Wensheng Yan[2], Duo Zhang[1]
[1]Xi'an Fanyi University, Xi'an 710105, Shaanxi, China
[2]Quanzhou University of Information Engineering, Quanzhou 362000, Fujian, China
Corresponding author: Duo Zhang (e-mail: liangdali_1999@163.com)

ABSTRACT With the continuous expansion of higher education, smart campuses increasingly rely on multi-source sensing, wireless communication, and real-time data propagation to support teaching, research, and engineering resource scheduling. However, the traditional static quota management model cannot dynamically respond to changes in teaching and research demand, resulting in structural contradictions between idle and scarce resources. Existing evaluation methods also remain weak in linking efficiency diagnosis with management intervention and cross-period optimization. To address these problems, this study constructs a higher education resource allocation and management framework integrating dynamic evaluation and intelligent optimization. Human, financial, material, and information resources are taken as four input factors, and a unified analytical view is established through multi-source data acquisition and preprocessing. A dynamic network SBM-DEA model is then designed to measure intertemporal allocation efficiency, while STL decomposition, spatial clustering, and LSTM prediction are used to extract spatiotemporal resource operation patterns and demand trends. Finally, a deep Q-network-based reinforcement learning algorithm is developed to form a closed loop from efficiency diagnosis to management intervention. The results show that the proposed model effectively identifies efficiency evolution trends and fairness differences among colleges. The median resource demand response lag decreased from 17.65 h in 2018 to 10 h in 2023, while the strategy adoption conversion rate increased from 0.551 to 0.861. The study provides a big-data-driven decision tool for precise resource allocation, and its sensing-data fusion logic is also relevant to engineering management scenarios involving electromagnetic signal acquisition, wireless monitoring, and distributed information propagation.

INDEX TERMS big data driven, resource allocation, dynamic network sbm-dea, wireless sensing, signal propagation

I. INTRODUCTION

As higher education scales and the demand for high-quality talent grows, the rational allocation of educational resources has become a key indicator for measuring the modernization of university governance and its ability to support advanced engineering research. In recent years, the rapid advancement of smart campus construction has enabled universities to collect and store real-time multi-source heterogeneous data from academic management, one card, online teaching platforms, and IoT devices, providing an unprecedented data foundation for accurately grasping the operation status of resources [1]. However, the current resource allocation in most universities still follows the traditional annual budget planning and static quota management model, which makes it difficult to dynamically respond to the real-time demand changes of teaching and research activities, leading to the increasingly prominent structural contradiction of idle and scarce resources [2-3]. How to use the perception ability of big data to break the static constraints of resource allocation and achieve the transformation from passive allocation to active adaptation has become a core challenge that urgently needs to be overcome in the field of higher education management. This study is based on the optimization of entire higher education resource allocation processes through big data. The specific research targets include 4 input factors (human, financial, material, and information resources) and three output dimensions (the quality of talent cultivation; effectiveness for innovation; and contribution to society) [4-5]. The process of resource allocation is vast and complex, with multiple levels and stage systems; it includes 2 primary functional links: support for teaching and converting achievements. There are obvious input-output time lags and cross period carry over characteristics between them [6- 7]. Meanwhile, the process of resource utilization is influenced by multiple factors such as semester cycles, course selection fluctuations, and research project progress, exhibiting typical spatiotemporal dynamics and non stationarity [8-9]. Therefore, not only should attention be paid to the rationality of static stock allocation, but more emphasis should be placed on dynamic diagnosis and precise regulation of resource flow processes, usage intensity, and demand adaptation capabilities [10]. Existing research has accumulated rich results in evaluating the efficiency of resource allocation in higher education. Early research often used data envelopment analysis to statically measure the efficiency of resource allocation at the university or college level, and calculated the relative efficiency score by constructing an input-output indicator system [11]. Yu Y et al. used a three-stage data envelopment analysis to evaluate the effectiveness of higher education fiscal expenditures in 31 provinces of China, in order to understand how the determinants of higher education financing efficiency vary over time and space [12]. Some scholars have introduced the DEA model to deconstruct the process of resource allocation in universities into sub stages of teaching support and achievement transformation, revealing the internal transformation bottleneck hidden by traditional single-stage efficiency evaluation [13-14]. Some studies have also adopted the SBM model, which overcomes the efficiency estimation bias caused by the neglect of non radial relaxation problems in traditional radial models by directly handling slack variables with input redundancy and output deficiency [15]. Sozen M et al. used slack variable-based measurement data envelopment analysis to evaluate the performance of some universities, providing practical suggestions for the effective use of resources [16]. In recent years, with the development of machine learning technology, algorithms such as LSTM and isolated forest have been applied to resource demand prediction and anomaly detection, providing new technical paths for dynamic management [17]. There still exist several clear limitations to current methods: evaluation and decision-making are disconnected, and thus the outcome of efficiency diagnostics cannot be immediately translated into a matching resource allocation plan; and since most models use cross-sectional data, they are incapable of adequately characterizing the long-term effects of intertemporal resource allocation and the mechanisms of dynamic carryover from one resource to another. Finally, the creation of optimization plans is not based on a system-level approach; therefore, they do not include multi-cycle and multi-objective collaborative optimization. Using an integrated dynamic evaluation and intelligent optimisation approach to building a management framework for resource allocation in higher education is the objective of this paper. This new method develops a closed-loop methodology for higher education resource allocation based on big data, allowing for the entire process of evaluation of efficient resource use through the ultimate optimisation of how resources are managed. This research study uses both multi-source data acquisition and preprocessing technology to develop a single data analysis view that captures the entire input to output process. Furthermore, it develops an efficiency measurement model that employs dynamic network SBM-DEA and incorporates carry-over variables and stage connection mechanisms to obtain precise evaluations of resource allocation efficiency across different time periods. Based on this framework, spatiotemporal features and demand patterns of resource usage can be determined through methods such as STL decomposition, spatial clustering analysis, and LSTM forecasting. Lastly, the development of a reinforcement learning-based optimization strategy generation algorithm using deep Q-networks is based on the efficiency evaluation results and demand forecasting as state inputs, thereby ensuring a feedback loop from efficiency assessment to management intervention. This research provides a systematic theoretical basis for transforming university resource allocation from experience-based to data-driven models, ensuring that management decision-making becomes scientific, accurate, and intelligent enough to support the complex requirements of fiber morphology and manufacturing excellence.

II. METHODS

A. MULTISOURCE EDUCATIONAL DATA ACQUISITION AND PREPROCESSING

Multi-source educational data acquisition relies on the university's information infrastructure. The architecture for multi-source educational big data acquisition and fusion is shown in Figure 1.

As shown in Figure 1, during the data storage and data cleaning and transformation stages, heterogeneous data access from the academic affairs management system, ID card system, online teaching platform, and Internet of Things (IoT) platform is achieved through a standardized API (Application Programming Interface) interface and direct connection to the database. To address the differences in data formats, a unified data collection platform is adopted. Apache Kafka is used to build a real-time data pipeline to ensure transmission stability and integrity under high concurrency scenarios [18]. For the missing value problem, a differentiated imputation strategy is adopted based on the missing value mechanism: for continuously variable variables with random missing values, a weighted imputation model is constructed based on the K-nearest neighbor algorithm [19], the core calculation of which is:

$$\hat{x}_{ip} = \frac{\sum_{k \in K_i} \omega_{ik} x_{kp}}{\sum_{k \in K_i} \omega_{ik}}, \quad \omega_{ik} = \frac{1}{d_{ik} + \epsilon} \tag{1}$$

In formula (1), $\hat{x}_{ip}$ represents the interpolated value of the p-th dimension feature of the ith sample, K_i is the set of K nearest neighbors of the ith sample, x_{kp} is the feature

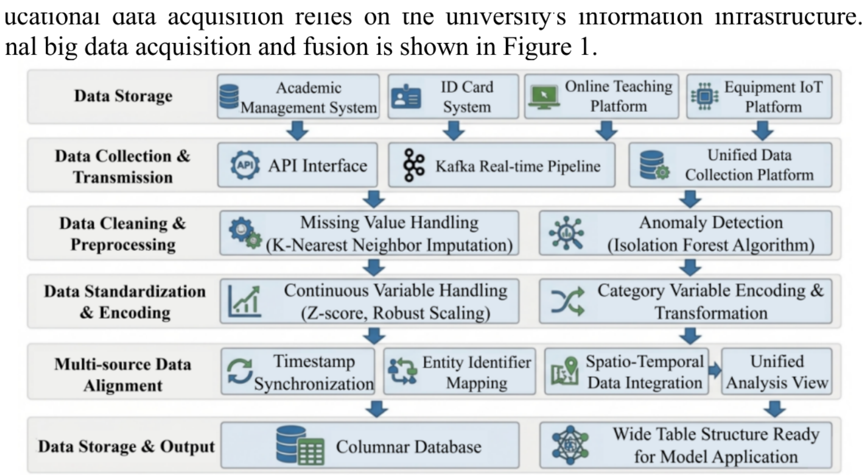

FIGURE 1. Multi-source educational big data collection and fusion architecture diagram

observation value of the neighboring samples, d_{ik} is the Mahalanobis distance-based metric between samples, and ϵ is the regularization constant to prevent division by zero. Outlier detection employs the isolated forest algorithm, which identifies outliers deviating from the normal distribution pattern by constructing a random segmentation tree. Records with equipment usage time exceeding the physical constraint threshold are marked and manually reviewed [20-21]. To eliminate the interference of differences in the dimensions of multi-source indicators on subsequent modeling, continuous variables are standardized. When adapting to the normal curve, data is transformed by Z-scores; when bias is present, Robust Scaling replaces the mean with the median and the Standard Deviation with the Interquartile Range to diminish the effect of outliers. One-hot Encoding and Target Encoding convert categorical data into numerical data for compatibility with algorithms. Multi-source data alignment is accomplished using timestamps and entity identifiers as a dual primary key in order to establish a mechanism for the spatio-temporal association between records. In the case of heterogeneous identifiers for an individual entity within different systems, an entity parsing approach that employs rule-based matching and fuzzy similarity methods is implemented to derive the appropriate attribution of records via the use of edit distance and jaccard coefficient. At the level of spatiotemporal alignment, a time window sliding strategy is introduced to aggregate minute level behavior logs into hourly granularity, and associate and fuse them with structured data such as course schedules and device reservation records. For cross system event sequences, a weight allocation method based on time decay is adopted to ensure that recent behavior occupies a higher priority in the analysis view. The final unified analysis view is stored in a column based database in a wide table structure, supporting the rapid retrieval and aggregation of high-dimensional features for subsequent efficiency evaluation models. The entire preprocessing process is managed through Airflow for task scheduling and dependency management, ensuring the traceability and reproducibility of the data pipeline. Building upon this robust data foundation, the research transitions from technical data engineering to management evaluation by performing a correlation diagnosis was conducted on various input-output indicators, and the results are shown in Table 1. There is a high correlation between the average teaching budget per student and the average equipment value per student, and it is necessary to perform dimensionality reduction in the subsequent indicator system construction to avoid the interference of information overlap on efficiency evaluation. The correlation matrix of input-output indicators is shown in Table 1.

Table 1 shows the Pearson correlation coefficient matrix of 8 core input-output indicators in higher education resource allocation, revealing significant differential correlation characteristics among the indicators. Among them, the correlation coefficient between per capita teaching funding and per capita equipment value reached 0.83, the per capita equipment value and digital resource access reached 0.79, and the per capita scientific research output and patent conversion amount were 0.81, all close to or exceeding the redundancy threshold of 0.8, indicating strong information overlap within the dimensions of financial, material investment, and scientific research innovation output; The correlation coefficient between the equivalent number of full-time teachers and the amount of patent conversion is only 0.38, and the correlation coefficient between the degree granting rate of graduates and the number of digital resource visits is 0.55, reflecting the relatively loose relationship between human resource investment and market-oriented achievement transformation, basic teaching resources and high-end output. This pattern of correlation differentiation suggests that if all indicators are directly included in the efficiency evaluation model, weight allocation may be distorted due to multicollinearity. Therefore, highly correlated indicators need to be screened or fused based on thresholds.

TABLE 1. Correlation matrix analysis of input output indicators

Indicator	Full-time Teacher Equivalent	Per-student Teaching Expenditure	Per-student Equipment Value	Digital Resource Access Volume	Graduate Degree Award Rate	High-quality Employment Rate	Research Output per Faculty	Patent Commercialization Amount
Full-time Teacher Equivalent	1	0.62	0.58	0.47	0.51	0.49	0.55	0.38
Per-student Teaching Expend.	0.62	1	0.83	0.71	0.67	0.63	0.72	0.54
Per-student Equipment Value	0.58	0.83	1	0.79	0.61	0.58	0.69	0.60
Digital Resource Access Volume	0.47	0.71	0.79	1	0.55	0.52	0.64	0.57
Graduate Degree Award Rate	0.51	0.67	0.61	0.55	1	0.76	0.70	0.63
High-quality Employment Rate	0.49	0.63	0.58	0.52	0.76	1	0.68	0.61
Research Output per Faculty	0.55	0.72	0.69	0.64	0.70	0.68	1	0.81
Patent Commercialization Amt	0.38	0.54	0.60	0.57	0.63	0.61	0.81	1

B. CONSTRUCTION OF INPUT OUTPUT INDICATOR SYSTEM FOR DYNAMIC EVALUATION

To eliminate the differences in metrics and information overlap between indicators, the original data is first standardized, and then the Pearson correlation coefficient matrix is calculated to identify multicollinearity. A correlation coefficient threshold of 0.8 is set, and indicators exceeding this threshold are considered to have information redundancy and require further screening. For highly correlated redundant indicators, principal component analysis is used for dimensionality reduction. The principal component extraction process is based on eigenvalue decomposition of the covariance matrix [22-23], which is mathematically expressed as:

$$\Sigma v_k = \lambda_k v_k, \quad \sum_{k=1}^{m} \lambda_k \geq 0.85 \sum_{j=1}^{p} \lambda_j \tag{2}$$

In formula (2), Σ is the covariance matrix of the standardized data, v_k is the eigenvector corresponding to the kth principal component, λ_k is the eigenvalue, m is the number of extracted principal components, and p is the total number of original indicators. Through this process, the originally strongly correlated average equipment value and laboratory area were merged into a comprehensive indicator of "material support capability", effectively reducing the complexity of the model and preserving the vast majority of the original information, solving the problem of indicator conflicts caused by different statistical calibers between multi-source data, and ensuring the independence of input variables. When determining the final indicator weights, the Analytic Hierarchy Process (AHP) combined with expert scoring mechanism is used to construct pairwise comparison judgment matrices [24-25]. Invite experts in the field of higher education management to evaluate the relative importance of indicators based on the Analytic Hierarchy Process, forming a positive and negative matrix. To ensure the logical consistency of expert judgments, the Analytic Hierarchy Process (AHP) method requires consistency testing of the judgment matrix. When the consistency indicators do not meet the requirements, the judgment needs to be adjusted [26-27], and the consistency ratio CR value needs to be calculated. The inspection formula is defined as:

$$CR = \frac{\lambda_{\max} - n}{(n-1) \cdot RI} \tag{3}$$

In formula (3), $\lambda_{\max}$ represents the maximum eigenvalue of the judgment matrix, n is the order of the matrix, and RI is the same order random consistency index. When $CR <$ 0.1, the judgment matrix has satisfactory consistency, otherwise it needs to be adjusted until it passes the test. By using the geometric mean method to aggregate the weight vectors of multiple experts, determine the combined weights of each secondary indicator, eliminate the interference of subjective preferences of a single expert on the evaluation results, and enhance the interpretability and expert-driven consistency of weight allocation. While AHP introduces expert judgment, it provides a structured qualitative-quantitative bridge that complements the purely data-driven DEA stage. The final indicator system consists of 7 criterion layers and 14 scheme layer indicators, each of which defines a clear data collection criteria and calculation logic. In addition to covering the static available resource allocation, this system includes the dynamic growth in resource levels for their supply, e.g., how often equipment is used and how quickly funds are expended to assess if an accurate assessment can be achieved to assess marginal resources used. The evaluation index system and data sources for the efficiency of higher education resource allocation are shown in Table 2.

The evaluation index system for the efficiency of higher education resource allocation constructed in Table 2 closely revolves around the research theme of "precise resource allocation and efficiency improvement driven by big data". It adopts a dual layer architecture design of "four-dimensional input three-dimensional output", covering four input elements of human resources, financial resources, material resources, and information resources, as well as three output dimensions of talent cultivation quality, scientific research innovation efficiency, and social service contribution. The input indicators, such as the equivalent number of full-time teachers, average teaching budget per student, average equipment value per student, and digital resource access volume, respectively depict the static stock and dynamic increment of resource allocation from the perspectives of teacher structure, funding guarantee, hardware support, and information technology level; Output side indicators such as degree granting rate, proportion of high-quality employment, per capita scientific research output, and patent conversion amount focus on the ultimate benefits and social value realization of educational activities. Each indicator has clearly defined the calculation criteria and data sources, relying on the integration of heterogeneous data from multiple systems such as academic affairs, finance, assets, and scientific research to ensure the traceability and empirical reliability of evaluation results. The design logic of this indicator system originates from the underlying mechanism requirements of the dynamic network SBM-DEA model: on the one hand, by using principal component analysis and

TABLE 2. Evaluation index system and data source table for efficiency of higher education resource allocation

Criteria Layer	Indicator Layer	Indicator Description	Data Source
Human Resource Input	Full-time Teacher Equivalent	Total teacher equivalent after weighting by professional title	Educational Administration System, Personnel System
Human Resource Input	Proportion of Senior Title Teachers	Proportion of teachers with senior professional titles among all full-time teachers	Personnel System
Financial Resource Input	Educational Expenditure per Student	Cumulative routine teaching expenditure per term / Full-time equivalent student enrollment	Financial System
Financial Resource Input	Scientific Research Expenditure	Quarterly research funds arrival rate	Scientific Research Management System
Physical Resource Input	Value of Teaching and Research Equipment per Student	Total value of equipment / Full-time equivalent student enrollment	Asset Management System
Physical Resource Input	Proportion of Usable Laboratory Area	Usable laboratory area / Total building area	Real Estate Management System
Information Resource Input	Digital Resource Access Volume	Monthly average visits to library resources to library electronic resources	Library System, Online Learning Platform
Information Resource Input	Online Course Offering Rate	Number of online courses offered / Total number of courses	Educational Administration System, Online Learning Platform
Talent Cultivation Output	Graduate Degree Award Rate	Number of students actually awarded degrees / Number of students expected to graduate	Educational Administration System, Student Affairs System
Talent Development Output	High-quality Employment Rate	Proportion of graduates employed in state-owned enterprises, government agencies, or pursuing further education	Career Guidance Center, Alumni System
Scientific Research Innovation Output	Scientific Research Output per Capital	Weighted score of papers, patents, research awards, etc. / Number of faculty members	Scientific Research Management System, CNKI, Patent Database
Scientific Research Innovation Output	Patent Commercialization Amount	Annual contract amount from patent commercialization	Technology Transfer Office, Financial System
Social Service Contribution Output	Horizontal Project Contribution	Research funds received from enterprise-entrusted projects	Scientific Research Management System
Social Service Contribution Output	Continuing Education Revenue	Total revenue from non-degree continuing education	Financial System, School of Continuing Education

correlation diagnosis to eliminate multicollinearity between indicators, the independence of input variables is guaranteed, and efficiency estimation bias caused by information overlap is avoided; On the other hand, the indicator system takes into account both static stock and dynamic flow characteristics, such as incorporating process indicators such as equipment usage frequency and fund execution rate, . To bridge the gap between high-frequency IoT/ID card data and the evaluation model, the system employs a multi-scale temporal aggregation strategy. Real-time behavior logs are aggregated into granular monthly or quarterly metrics—such as usage intensity and fund execution velocity—which are then fed into the dynamic network SBM-DEA to ensure that the low-frequency evaluation remains sensitive to real-time operational shifts. The multi-source heterogeneity of data sources supports the state space construction of STL-LSTM time series prediction and reinforcement learning strategy generation algorithms, enabling the efficiency evaluation results to be deeply integrated with resource demand trends and spatiotemporal operational characteristics, forming a closed-loop management chain of "diagnosis prediction optimization". Ultimately, this indicator system not only provides quantifiable and comparable evaluation benchmarks for the allocation of higher education resources, but also promotes the transformation of management decision-making from empirical judgment to intelligent adaptation through a data-driven paradigm, achieving collaborative optimization of efficiency improvement and fairness protection.

C. EFFICIENCY EVALUATION MODEL BASED ON DYNAMIC NETWORK SBM-DEA

This study defines every college as an individual decision-making unit. The production possibility set consists of two stages for providing teaching support and transforming teaching to student achievement. In doing so, this research reveals how colleges allocate resources internally through the "black box" that each college contains. Stage 1 is about converting input into an output, which involves converting resource expenditures (such as teacher and student costs) into teaching activities (standardized workload, curriculum quality index). Stage 2 is converting teaching activities to provide output. Intermediate outputs of stage 1 are used as inputs for stage 2, while final outputs are based on the graduate degree granting rate and scientific research output per capita. The intermediate outputs connect the two stages and therefore should be maximized in relation to the outputs produced during the first stage of activity and used to minimize the inputs needed to achieve the desired outputs of the second stage of activity. This constraint on the dualrole structure ensures coherence of the internal process. In order to describe the characteristics of cross periodic resource allocation, the model describes a carry over variable from period-to-period. The unused capacity of equipment, and the surplus funds not used in the prior cycle are identified as free carry-over variables. These carry over variables can continue to be used as free carry-overs in the next cycle. This overcomes the restrictions of conventional DEA models that only measure one static snapshot at a time and achieves comprehensive cycle efficiency tracking [28-29]. Thus the dual-role structure is an effective solution for addressing the problem of the delay time lag between resource input and output, ensuring that the evaluative results reflect the long term educationally related laws of postsecondary education. The kernel density estimation of the efficiency value is depicted in Figure 2.

As illustrated by Figure 2, higher education resource allocation efficiency values display a probability density distribution characteristic for overall efficiency as well as first-stage efficiency and second-stage efficiency via three kernel density estimate curves. The three curves show a clear hierarchical distribution pattern, reflecting the efficiency differences and underlying mechanisms at different stages of resource allocation. The peak position of the efficiency curve in the second stage is relatively left, indicating that the overall efficiency level of the achievement transformation stage is low. This reveals that universities have obvious bottlenecks in transforming teaching activities into the final quality of talent cultivation and scientific research innovation achievements. The distribution of the efficiency curve in the first stage is centered, indicating that the efficiency of converting resource investment into teaching activities is at a moderate level, reflecting that universities have a

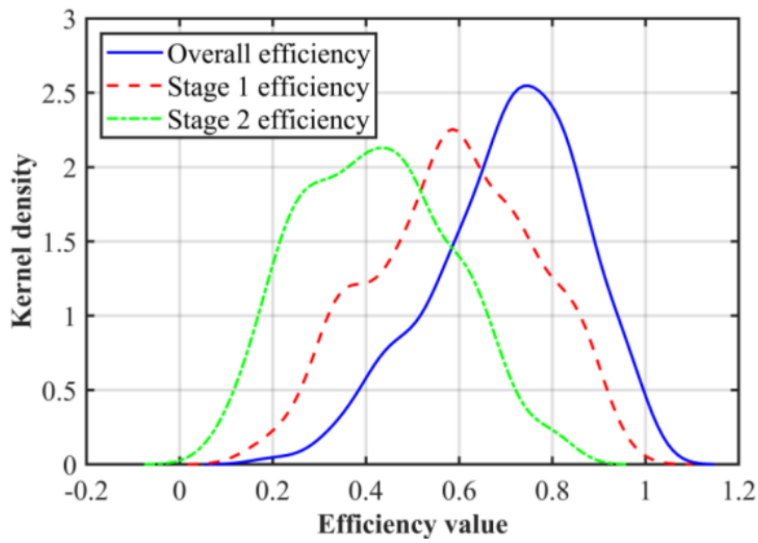

FIGURE 2. Efficiency value kernel density estimation

certain foundation but still have room for optimization in the allocation of teaching resources. The peak position of the overall efficiency curve is closest to the right and the distribution is relatively concentrated. The mechanism behind this phenomenon is that the dynamic network SBM-DEA model effectively characterizes the characteristics of cross period resource allocation by introducing carry over variables and stage connection constraints, enabling the overall efficiency evaluation to comprehensively consider the synergistic effects of the two stages. The morphological differences in nuclear density distribution also reflect the degree of dispersion of efficiency values at different stages. The relatively dispersed efficiency distribution in the second stage indicates significant heterogeneity in the ability of each decision-making unit to transform results, which is closely related to the differences in disciplinary characteristics, research foundations, and management levels of universities. This distribution feature provides important state space input basis for the formulation of subsequent reinforcement learning optimization strategies. The model adopts a non radial SBM directional distance function to directly handle input redundancy and output deficiency slack variables, overcoming the shortcomings of traditional radial models that ignore slack variables [30-31]. The objective function aims to minimize the slack ratio between input and carry over variables, while maximizing the slack ratio of output. Its mathematical expression is:

$$\rho^* = \min \frac{1}{T}\sum_{t=1}^{T} w^t \left[1 - \frac{1}{m+n_{carry}}\left(\sum_{i=1}^{m}\frac{\bar{s}_{it}^{-}}{x_{i0t}} + \sum_{k=1}^{n_{carry}}\frac{s_{kt}^{carry}}{z_{k0,t,t+1}}\right)\right] \tag{4}$$

In formula (4), ρ^* represents the overall efficiency value, T is the total number of evaluation cycles, w^t is the weight coefficient of the t-th period, m is the number of input indicators, n_{carry} is the number of carry over variables, $\bar{s}_{it}^{-}$ is the relaxation variable of the ith input in the tth period, x_{i0t} is the observation value of the ith input of the decision unit in the tth period, s_{kt}^{carry} is the relaxation value of the kth carry over variable, and $z_{k0,t,t+1}$ are the observation values carried over from the tth period to the $t+1$th period. To ensure that production possibilities are logically consistent and that virtual output does not inflate efficiency values, the inter stage connections and carry-over from one time period to another need to be constrained.

The method of solving uses a linear programming algorithm to iteratively compute the efficiency values for each decision unit within each timeframe. By separating out the objective function into three dimensions, complete efficiency results can be retrieved in terms of overall efficiency, stage efficiency, and period efficiency, allowing for the isolation of the source of inefficiency from resource input redundancy, obstacles between stages of transformation, and cross-period configuration imbalance. The results generated from the model for efficiency scores are also used to make horizontal comparisons with respect to the amount and type of resources allocated to various colleges and serve as input to the state space for automated reinforcement learning optimization algorithms, thereby providing management optimization strategies based on observed efficiency bottlenecks as opposed to intuitive estimates.

D. RESOURCE SPATIOTEMPORAL OPERATION FEATURE EXTRACTION

Due to the changing nature of the resource consumption time series data, the STL decomposition technique was employed to separate the time series into three separate sections: trend, seasonal and residual. Using the weekly classroom utilization rates as an example, the seasonal window parameters are set to align with an annual cycle model, whereas the trend window parameters are readily modified to fit local weighting regression criteria. The decomposition procedure is accomplished with local weight least squares incremental optimization, whose objective function is represented as:

$$\min_{\beta}\sum_{i=1}^{n}K\left(\frac{t-t_i}{h}\right)\left(y_i-\mathbf{x}_i^T\beta\right)^2+\lambda\int\left[f''(t)\right]^2\,dt \tag{5}$$

In formula (5), $K(\cdot)$ is the cubic kernel function, h is the bandwidth parameter, y_i is the observed value, $\mathbf{x}_i$ is the basis function vector, β is the estimated coefficient, λ is the smoothing parameter, and $f''(t)$ is the second derivative of the trend function. By separating seasonal fluctuations from long-term trends, residual sequences can effectively identify abnormal occupancy events and provide temporal benchmarks for resource scheduling. Spatial clustering analysis generates a multidimensional feature vector using geographic coordinates, floor heights, and room use codes for a campus. It uses Euclidean distance to measure how close the spatially clustered area is to each other. For the gridded spatial units defined by rules, a K-means algorithm is used to cluster the resource loading levels. The correct number of clusters is found by combining an elbow method with the contour coefficients of high loading, balanced, and not loaded spatial areas across [32-33]. For irregularly spaced spatial point clusters, like those found in laboratories, a density-based spatial clustering algorithm (DBSCAN) based on density reachability is used to establish clusters and their structure. The radius of a neighbouring area is found from a k-distance and is defined in terms of twice the feature dimensions for the minimum number of samples (MinPts) [34]. The predicted requirements for resources is created from a series learning approaches using an LSTM architecture. The model uses the following input features: historic rates of usage by way of each enrolled student; each research project; and semester dummy variables. The output of the model is a probability distribution of what the expected demands for resources can be in the next week, where quantile regressions are used to create confidence intervals that represent risk of allocating resources flexibly.

E. RESOURCE ALLOCATION OPTIMIZATION STRATEGY GENERATION ALGORITHM BASED ON EFFICIENCY DIAGNOSIS AND DEMAND PREDICTION

This paper develops a flexible system of allocating resources through Markov decision processes by changing the multi-cycle resource allocation problems into sequential decision optimization problems. The S_t state space consists of DEA efficiency evaluation values from the tth period, the real time usage vectors of different resources, and the future demand characteristics as estimated by an LSTM, thus creating a state representation that has both historical performance information as well as future trend information. To expedite the convergence of the neural net, every input can be Z-scored. The Action Space A_t is comprised of a finite set of instructions for allocating resources among resource allocation sources; The actions consist of reallocating flexible operational funds within institutional limits, classroom reallocation plan, and equipment sharing permission adjustment. To ensure practical feasibility, the budget adjustment space is constrained by pre-defined administrative quotas, rather than allowing for unconstrained algorithm-led changes. A composite reward function R_t is created to direct intelligent agents in learning allocation strategies that result in an equal balance of efficiency and fairness, which can be mathematically expressed as:

$$R_t=\eta_1(\rho_{t+1}-\rho_t)+\eta_2(G_{base}-G_{t+1})-\eta_3\|\Delta r_t\|_1 \tag{6}$$

In formula (6), ρ_t is the DEA efficiency value for the tth period, G_t is the Gini coefficient for resource allocation, Δr_t is the resource adjustment vector, and η_1,η_2,η_3 are the weight coefficients for efficiency gain, fairness improvement, and adjustment cost, respectively. It provides efficiency and effective distribution of wealth through participation incentives, minimizes large or frequent variations in available resources through disincentives, maintains consistency and stability of a management plan, and avoids creating a chaotic situation in the management system due to excessive variations in resource availability. Adopting the deep Q-network algorithm for strategy training, a neural network structure is constructed that includes a state input layer, a double-layer hidden layer, and an action value output layer. The network parameters are iteratively updated by minimizing the temporal differential error, and the loss function is defined as the mean square error between the predicted Q value and the target Q value:

$$\begin{aligned}L(\theta)=E_{(s,a,r,s')\sim D}\Big[&(r+\gamma\max_{a'}Q(s',a';\theta^-)\\&-Q(s,a;\theta))^2\Big]\end{aligned} \tag{7}$$

In formula (7), θ is the current network parameter, θ^- is the target network parameter, γ is the discount factor, and D is the experience replay buffer. Introducing an experience replay mechanism to break the temporal correlation of data and utilizing a fixed Q-value target in the target network to improve training convergence. During the training process, a ϵ - growth strategy is adopted to balance exploration and utilization. In the initial stage, the random exploration probability is increased to cover the state space, and gradually decreases to the threshold as the training rounds increase, ensuring that the agent can discover the global optimal configuration path rather than getting stuck in local optima [35]. The model is iteratively trained in the simulation environment until the cumulative reward value converges, indicating that the policy network has mastered the dynamic adaptation law of resources. The loss curve is shown in Figure 3.

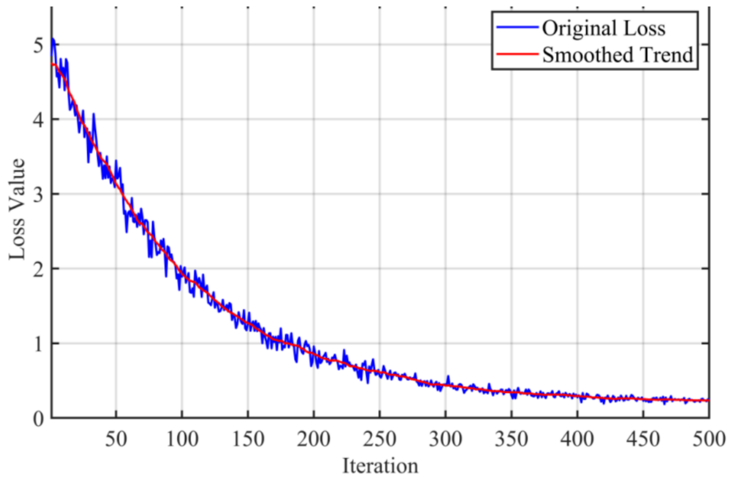

FIGURE 3. Training loss curve

Figure 3 shows the evolution curve of the training loss function of the deep Q-network in the process of generating optimization strategies for higher education resource allocation. From the overall trend, the loss value shows a continuous downward trend with the increase of iteration times, gradually converging from the initial high level. This reflects the continuous optimization and improvement of the policy network by the reinforcement learning agent during the training process. Although there is a certain degree of fluctuation in the original loss curve, the overall downward trend is clear. This volatility is due to the data diversity brought by random sampling in the experience replay mechanism, as well as the random action selection of the ϵ - growth exploration strategy in the early stages of training. The smooth trend line reveals the inherent laws of model learning more clearly, and its monotonically decreasing feature indicates that the temporal difference error is continuously decreasing, and the deviation between the predicted Q value and the target Q value is gradually decreasing. During the optimization plan generation phase, real-time monitoring data is input into the trained strategy network to output the optimal probability distribution of each decision unit's actions. Select the action with the highest probability as the recommended configuration plan, and combine it with administrative constraints for feasibility verification, such as budget limits and physical space limitations. This algorithm implements a closed loop from efficiency diagnosis to management intervention, solving the problems of traditional resource allocation relying on manual experience and response lag, supporting management to make data-driven and accurate decisions, and ensuring that resources flow to high-performance and high demand areas.

III. EXPERIMENTS AND VALIDATION

A. EXPERIMENTAL DESIGN

This study selects a certain university as an empirical case. The university has a complete range of disciplines, covering ten major colleges including liberal arts, science, engineering, and medicine, and has typical characteristics of higher education resource allocation and complex management scenarios. The sample selection meets the criteria of being representative and generalizable across disciplines to prevent model bias related to individual disciplinary contexts that show high dependence of the disciplines of science and engineering on equipment, while humanities have sensitivities to funding in different ways. Throughout data collection, this research upholds the requirement of confidentiality protection by using only non-identifiable aggregated indicators and anonymised identifiers for statistical purposes from teachers' and student's sensitive personal data. This research has successfully developed a complete chain panel dataset by connecting multiple different systems (academic affairs, finance, and assets) to aggregate resource inputs, process operations, and result outputs to ensure the experimental data reflects the dynamic evolutionary nature of resource allocation, and, therefore, solves the issue of one dimensional evaluations created by information silos. The data collection phase was planned to occur between 2014 and 2023, covering a period of twenty semesters; it represents an entire cycle of resource allocation adjustment during standard periodic teachers' academic years as well as during certain events outside those normal cycles, providing adequate basis for evaluating the continued stability of the actual model among various environments. The data collection content not only includes static resource stock data at the beginning and end of the semester, such as teacher quantity, budget, equipment assets, etc., but also includes dynamic operational data during the semester, such as student grade distribution, research project progress, classroom usage logs, etc. This high-frequency panel data structure solves the problem of traditional cross-sectional data being unable to capture time series dependencies, providing necessary temporal input basis for dynamic network SBM models and LSTM prediction models, ensuring that efficiency evaluation can reflect the cross period carry over effects of resource utilization. To verify the effectiveness and generalization ability of the model, the dataset is divided into a training set and a testing set. The data from eight semesters from 2014 to 2017 were used as the training set for model parameter learning and policy network training, and were not included in the effectiveness validation; The data from the twelve semesters from 2018 to 2023 can be used as the test set for effectiveness validation and error analysis. This time-based partitioning method strictly follows causal logic, avoiding the problem of inflated evaluation caused by future information leakage, and ensuring that the model learns historical patterns rather than data noise.

B. MEASUREMENT AND EVOLUTION TREND OF RESOURCE ALLOCATION EFFICIENCY

The comprehensive efficiency values of each college are calculated based on the dynamic network SBM-DEA model. The model considers each college as an independent decision-making unit, with human, financial, material, and information resources as input factors, talent cultivation quality, research and innovation efficiency, and social service contribution as output dimensions, and introduces carry over variables to characterize the characteristics of cross period resource allocation. By solving the objective function that includes input slack variables and output slack variables, the comprehensive efficiency values of each college from 2018 to 2023 are output, which comprehensively reflect the overall allocation efficiency of resources in the two stages of teaching support and achievement transformation. Based on the above calculation results, the dynamic evolution trend of the comprehensive efficiency values of each college is shown in Figure 4.

Figure 4 shows the dynamic evolution trend of the comprehensive efficiency values of ten colleges from 2018 to 2023, and the data comparison reveals a significant pattern of differentiated development. Colleges 1, 2, 3, and 6 are showing a continuous downward trend, with College 1 decreasing year by year from a high of 0.915 in 2018 and stabilizing at a low level of 0.4 in 2022-2023; College 2 also decreased from 0.853 to 0.4, while College 3 and College 6 converged from 0.679 and 0.628 respectively to around 0.4. The steady improvement trajectory of colleges 4, 8, and 10, with college 4 achieving a leapfrog growth from 0.4 to 0.842, college 8 rising from 0.543 to 0.95, and college 10 rising from 0.602 to 0.869. Colleges 1, 2, 3, and 6 with continuously declining efficiency values may face redundant resource investment in the teaching support stage or stage transformation obstacles, where investments such as equivalent full-time teachers and per student teaching funds cannot be effectively converted into intermediate outputs such as standardized teaching workload and course quality index; There may be bottlenecks in the transformation of intermediate outputs into final benefits such as degree granting rates for graduates and per capita scientific research output during the achievement transformation stage. On the contrary, the efficiency improvement trajectories of colleges 4, 8, and 10 indicate that they may have broken the shackles of traditional static quota management through their big data perception capabilities. With the help of multi-source data collection and spatiotemporal feature extraction, they have achieved precise diagnosis and dynamic regulation of the resource flow process, allowing resources to flow towards high-performance and high demand areas. The marginal benefit measurement of the student teacher ratio adopts a panel data regression model, and constructs an econometric analysis framework with the college semester as the observation unit. When interpreting marginal benefits, it is important to focus on the sign direction of the regression coefficient. A significant negative direction indicates that the expansion of the student teacher ratio has a suppressive effect on the quality of education, while a non significant positive direction suggests that teacher allocation has not yet reached the threshold of economies of scale. The measurement of research funding output elasticity is based on the form of a transcendental logarithmic production function, which performs logarithmic transformation on the original variables to eliminate heteroscedasticity and achieve direct estimation of elasticity. The scientific research achievements of the dependent variable are synthesized through multidimensional indicators, covering paper publication, patent authorization, and award acquisition. The entropy weight method is used to determine the weights of each sub indicator to avoid subjective weighting bias. The indicator results of the input-output efficiency of teacher resources and funding are shown in Figure 5.

Figure 5 shows the two-dimensional measurement results of the input-output benefits of faculty and funding. Figure 5 (a) shows that only 4 out of 10 colleges have a positive marginal contribution in the student teacher ratio marginal benefit data. College 7 ranks first with a marginal benefit value of 0.0909, while College 2 and College 8 have reached 0.0796 and 0.0718, respectively. This indicates that the current faculty configuration of these colleges has not yet reached the threshold of economies of scale, and moderately increasing the student teacher ratio can still have a positive impact on the quality of talent cultivation; The marginal benefits of the remaining six colleges are negative, with college 4 as low as -0.1825 and college 6 as -0.1576, indicating that the expansion of the student teacher ratio has had a suppressive effect on the quality of education. Figure 5 (b) shows that the elasticity of research funding output exhibits differentiated response characteristics, with elasticity coefficients ranging from 0.3478 to 0.7805. College 3 leads significantly with a high elasticity value of 0.7805, while College 7 and College 6 reach 0.699 and 0.6841, respectively, reflecting the high sensitivity of research output to funding investment in these colleges; In contrast, the elasticity of College 2 is only 0.3478, while College 1 and College 4 are about 0.43, indicating that the marginal conversion efficiency of their funding investment is relatively limited and needs to be optimized in conjunction with human capital synergy. The difference in research funding elasticity is identified through the estimation of bidirectional fixed effects beyond logarithmic production

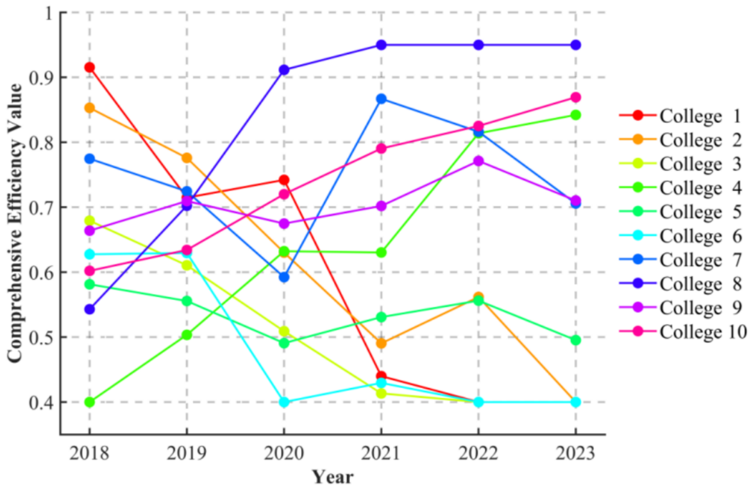

FIGURE 4. Trends in the comprehensive efficiency values of each college

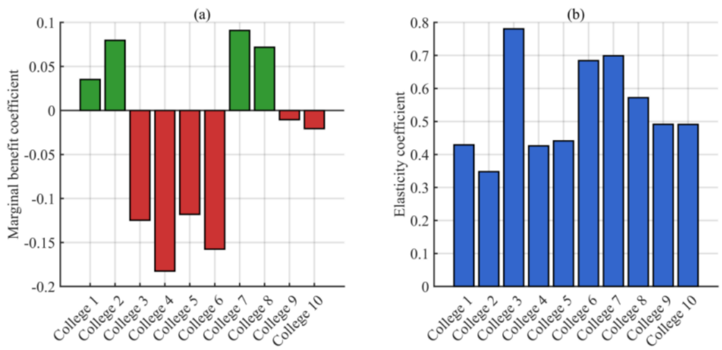

FIGURE 5. Efficiency indicators of teacher and fund input output
Figure 5 (a) Marginal benefit of student teacher ratio
Figure 5 (b) Elasticity of research funding output.

functions. High elasticity colleges are mostly in the capital driven innovation stage, and the synergistic effect between equipment platforms and research funding is significant. Low elasticity colleges rely more on human capital accumulation, and simply increasing funding is difficult to quickly convert into research output.

C. FAIRNESS IN RESOURCE ALLOCATION

The Gini coefficient measurement of per capita resources is based on colleges as the basic accounting unit. Firstly, the per capita funding values of each college are calculated and sorted in ascending order, and the cumulative quantity percentage and cumulative funding proportion of each college are synchronously calculated. This indicator dynamically tracks the trend of Gini coefficient changes over the years. The Theil index measurement of inter professional resources adopts a nested grouping strategy to calculate the degree of deviation between the average resources per student within each college and the overall average. The Gini coefficient and Theil index of internal resource allocation in universities are shown in Figure 6.

Figure 6 shows a dual dimensional evaluation framework for the fairness of internal resource allocation in universities, where Figure 6 (a) shows the Gini coefficient of per capita resources within the institution reflecting the overall distribution balance, and Figure 6 (b) reveals the structural characteristics of the sources of differences in the Theil index between various projects within the institution. The Gini coefficient of College 7 increased from 0.44 to 0.58 and reached 0.61 in 2022, indicating a continued intensification of the Matthew effect on resources; College 4 has been optimized from 0.41 to 0.28, reflecting the effective intervention of the balanced intervention mechanism; College 6 rebounded to 0.37 in 2023 after dropping to a low of 0.16 in 2022, reflecting the fragility of fairness under external shocks. The average Gini coefficient of 10 universities has increased from 0.37 in 2018 to 0.46 in 2023. In terms of Theil index, College 7 has consistently maintained a high level (0.60-0.71), indicating that structural inequality between disciplinary categories is the main contradiction; College 5 has jumped from 0.16 to 0.48, indicating a difference in management efficiency within the college; College 1 fell back to 0.36 amidst fluctuations, reflecting the synergistic convergence of inter group and intra group

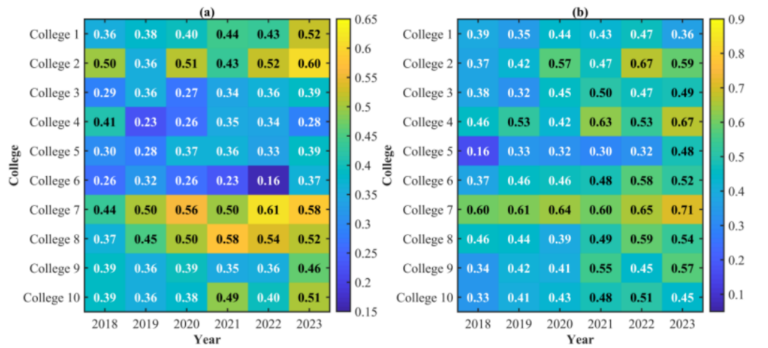

FIGURE 6. Equity index of resource allocation within universities
Figure 6 (a) Gini coefficient of per capita resources per student in the hospital
Figure 6 (b) Theil index between various projects within the hospital

differences. Cross analysis of two indicators shows that some colleges, such as College 2, have seen a synchronous increase in Gini coefficient and Theil index, indicating that the deterioration of fairness has both overall and structural characteristics. However, College 4 has improved its Gini coefficient but its Theil index has risen to 0.67, reflecting that there is still room for optimization in internal project allocation under overall balance.

D. TIME DELAY IN RESOURCE DEMAND RESPONSE AND ELASTICITY COEFFICIENT OF RESOURCE ALLOCATION

The delay measurement of resource demand response relies on event driven architecture to achieve accurate matching between demand signals and allocation actions. To begin, this article elaborates on the requirement-initiating principle, in which if either the maximum length of the classroom reservation application queue is exceeded or the number of equipment failure repair reports increases excessively or the number of students' preferred courses is outside the confidence interval this system automatically creates a demand event record with a timestamp. Second, the article follows up with a record of resource allocation management from the time demand events are generated to the point of resource allocation and execution. Finally, the article documents the resource allocation activities and calculate the total time taken from the initial demand event to the final execution. In cases where multiple demand events occur simultaneously, the priority queue can be used to arrange the demand events based on priority to prevent low-priority events from denying higher-priority responses. Resource allocation elasticities are calculated using a ratio-based method that takes into account the month-to-month change in resource allocation; as well as the relative quantities of resources available to students when compared based on semester-level granularity. The rate at which total resources are changing may be derived from various core performance measures like budget, faculty establishment, and classroom area, and adjusted each season to remove cyclical fluctuations (i.e. seasonal factors that occur over the course of an academic year). The rate of change of total students is calculated based on an equivalent number of students; and further aggregated by weighting to distinguish between the various levels of types of education (i.e. undergraduate, graduate and PhD) that exist in the education system. When performing the calculations for these two indices, a moving average filter was used to eliminate noise from the short term period before developing the elastic coefficients based on normalized change from anomalous fluctuations in a single semester. Figure 7 illustrates the annual variations in dynamic adaptability index.

Figure 7 depicts how well universities are adapting to changing resource allocation practices (2018-2023). Trends in two areas are available: delay in responding to demand for resources and the elasticity of resource allocation. The responses to both dimensions exhibit downward trends. With respect to delay in response to resource requests, the median value changed from 17.65 hours in 2018 to 10 hours in 2023; with respect to maximum and minimum values, they also declined from 20.1 and 13.9 hours to 12.6 and 5.6, respectively, which indicates greater agility with which universities can allocate resources and better control of extreme delays. With respect to the elasticity coefficient, the median changed from 1.12 in 2018 to 1.08 in 2023; also the difference between maximum and minimum values changed from 0.32 in 2018 to 0.30 in 2023. Hence, over the past six years, colleges have optimized both their speed to respond to requests for resources and their accuracy in regulating the use of available resources. Consequently, they have improved the dynamic adaptability of their resource management functions.

E. STRATEGY ADOPTION CONVERSION RATE AND RESOURCE ADJUSTMENT HIT RATE

The conversion rate of strategy adoption is tracked exclusively through the decision support system through the use of the logging/control module, which provides a means to capture the closed loop of suggestion generation and management feedback. Each time the decision support system generates resource configuration optimization suggestion content, the system automatically creates a structured record of the suggestion content, applicable objects, expected ben-

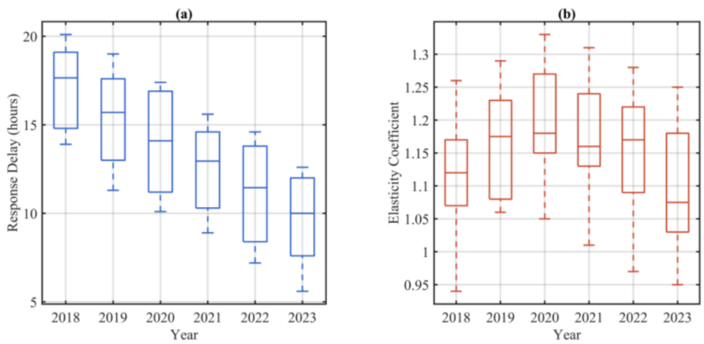

FIGURE 7. Annual change of dynamic adaptability index
Figure 7 (a) Annual distribution of resource demand response time delay
Figure 7 (b) Annual distribution of resource allocation elasticity coefficient

efits and time stamps that can be forwarded electronically to the management department's business terminal. The management department can then make their determination about adopting, partially adopting or rejecting each suggestion and can use the integrated feedback function of the system to provide a statement supporting the basis of their decision. The semester is defined and used for calculating the conversion rate and serves as the statistical time frame. The numerator is defined as the number of adopted or partially adopted suggestions, while the denominator is the total number of validly-generated suggestions provided by the system during that same period-except for any invalid or auto-filtered outputs due to missing data or model confidence that is below threshold. In some cases of adoption, a weighting factor of 0.5 is applied as a retention ratio to reduce the loss of accuracy resulting from the binary counting method used to compute the number of successes in terms of suggesting core elements. The hit rate of resource adjustment is measured using spatial set similarity analysis method, which evaluates the accuracy of resource allocation based on the physical space of the campus as the mapping carrier. The adjusted resource hotspot area is defined as a set of spatial units for incremental resource allocation or functional redistribution in the system optimization plan. The actual demand hotspot area is composed of spatial units with utilization rates exceeding the preset threshold based on the monitoring data of resource usage intensity at the end of the semester. The calculation of set similarity adopts Jaccard index, where the numerator counts the number of intersecting elements between two sets, that is, the spatial units predicted by the system to have high demand but actually showing high load, and the denominator counts the number of merging elements between the two sets, covering all spatial units recognized or actually having high demand by the system. To address the issue of inconsistent granularity of spatial units, geographic coding mapping is used to unify spatial identifiers from different sources into a standard room numbering system, ensuring the accuracy of set operations. The closer the hit rate is to 1, the higher the spatial distribution between the system's predictions and actual needs, and the stronger the foresight and accuracy of resource allocation. The adoption conversion rate and resource adjustment hit rate are shown in Figure 8.

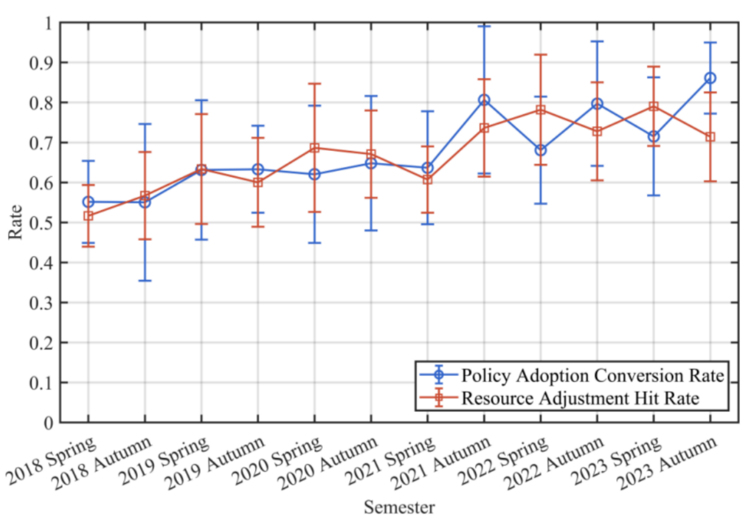

FIGURE 8. Data driven decision effectiveness indicators

The data-driven decision effectiveness indicators presented in Figure 8 include two core dimensions: strategy adoption conversion rate and resource adjustment hit rate. Data tracking from spring 2018 through fall 2023 demonstrates that while both indicators show a cumulative upward trending movement and differences in their evolutionary paths. The conversion rate of adopting strategies rose from 0.551 in spring 2018 to 0.861 in fall 2023. The success rate for adjusting resources showed an increase from 0.517 in spring 2018 to 0.791 in spring 2023 before dropping to 0.714 in fall 2023. From a dispersion standpoint, the standard deviation of the adoption rate reached 0.089 in fall 2023. This shows that the various colleges are solidifying their acceptance of suggested optimized strategies; however, the standard deviation of success rates ranged from 0.08 to 0.16, indicating that there is still a degree of uncertainty when predicting success based on spatial predictors, due to the cycles of semesters and other external factors.

IV. CONCLUSION

This article constructs a big data-driven framework for the allocation and management of higher education resources, facilitating the precise digital transformation of experimental research in fiber morphology and engineering. This article uses the dynamic network SBM-DEA model to measure cross period efficiency, combines STL decomposition and LSTM to predict demand trends, and generates optimization strategies through a deep Q-network to achieve a closed-loop system from efficiency diagnosis to management intervention. Empirical evidence shows that the model can effectively identify the differences in efficiency evolution and fairness of colleges, reduce the delay in resource demand response to 10 hours, and increase the strategy adoption conversion rate from 0.551 to 0.861. The research promotes the transformation of university governance from experiential to evidence-based, providing precision decision-making tools for resource allocation and the digital modernization of engineering research.

AUTHOR CONTRIBUTIONS

Conceptualization – Donghe Liang, Xiaoying Dong, Chan Bai, Wensheng Yan and Duo Zhang; methodology – Donghe Liang, Chan Bai, Wensheng Yan and Duo Zhang – Donghe Liang, Xiaoying Dong, Chan Bai, Wensheng Yan and Duo Zhang; writing-original draft preparation –Donghe Liang, Xiaoying Dong, Chan Bai, Wensheng Yan and Duo Zhang. 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] K. Polin, T. Yigitcanlar, M. Limb, and T. Washington, “The making of smart campus: A review and conceptual framework,” Buildings, vol. 13, no. 4, pp. 891-898, 2023, doi: 10.3390/buildings13040891.

[2] P. Ferrero L G and M. Salles-Filho S L, “Planning and resource allocation models in research-intensive universities: budget allocation and the search for excellence,” Humanities and Social Sciences Communications, vol. 12, no. 1, pp. 1-15, 2025, doi: 10.1057/s41599-025-04778-z.

[3] T. Zhu T, F. Yang, and J. Zhang Y, “Evaluating the efficiency of online course resource allocation in universities of China,” Evaluation and Program Planning, vol. 107, no. 1, pp. 102481-102486, 2024, doi: 10.1016/j.evalprogplan.2024.102481.

[4] W. Guo, “Improving higher education resource allocation efficiency and its spatial correlation for sustainable development in China,” Decision Making: Applications in Management and Engineering, vol. 7, no. 2, pp. 591-607. http://www.dmame-journal.org/index.php/dmame/article/view/1321, 2024.

[5] H. Guo, “The efficiency of resource allocation in higher education in india based on dea-tobit model,” Applied Mathematics and Nonlinear Sciences, vol. 8, no. 2, pp. 1837-1850. https://dialnet.unirioja.es/servlet/articulo?codigo=9264316, 2023.

[6] Z. Ye, R. Khanal, and Z. Cao, “Studying on the efficiency of higher education resource allocation and its influencing factors in the western China using DEA-Malmquist and Tobit models,” PLoS One, vol. 20, no. 10, Art. no. e0334090, 2025, doi: 10.1371/journal.pone.0334090.

[7] B. Chen, Y. Chen, Y. Sun, Y. Tong, and L. Liu, “The measurement, level, and influence of resource allocation efficiency in universities: empirical evidence from 13 “double first class” universities in China,” Humanities and social sciences communications, vol. 11, no. 1, pp. 1-16, 2024, doi: 10.1057/s41599-024-03461-z.

[8] H. Wang and Y. Wei, “Spatial dynamics and temporal trends of higher education institutions in China: a visual analysis,” Asia Pacific Journal of Education, vol. 1, no. 1, pp. 1-23, 2025, doi: 10.1080/02188791.2025.2568483.

[9] Y. Xu, “Research on Optimized Allocation Model of Educational Resources in Colleges and Universities Based on Big Data,” Applied Mathematics and Nonlinear Sciences, vol. 9, no. 1, pp. 1-7. https://elibrary.ru/item.asp?id=80006080, 2024.

[10] Z. Xu and Z. Xu, “Research on circulation mechanism of digital course resources from the perspective of information ecology theory,” Journal of Cases on Information Technology (JCIT), vol. 26, no. 1, pp. 1-21, 2024, doi: 10.4018/JCIT.335949.

[11] Z. Liu, H. Xiong, and Y. Sun, “Will online MOOCs improve the efficiency of Chinese higher education institutions? An empirical study based on DEA,” Sustainability, vol. 15, no. 7, pp. 5970-5982, 2023, doi: 10.3390/su15075970.

[12] Y. Yu, Y. Tingting, L. Ruoxi, and W. Xinxin, “Effectiveness of higher education financing: DEA and SFA modelling,” Business Ethics and Leadership, vol. 8, no. 1, pp. 172-189, 2024, doi: 10.61093/bel.8(1).172-189.2024.

[13] Y. Chen, Y. Pan, H. Liu, H. Wu, and G. Deng, “Efficiency analysis of Chinese universities with shared inputs: An aggregated two-stage network DEA approach,” Socio-Economic Planning Sciences, vol. 90, no. 1, pp. 101728-101736, 2023, doi: 10.1016/j.seps.2023.101728.

[14] Q. Zhong, “Allocation Efficiency of Higher Education Resources in China: A Three-Stage DEA Model-Based Calculation,” Frontiers in Economics and Management, vol. 6, no. 8, pp. 51-67, 2025, doi: 10.6981/FEM.202508_6(8).0006.

[15] J. Wang, W. Zhang, M. Zhao, X. Lai, L. Chang, and Z. Wang, “Efficiency of higher education financial resource allocation from the perspective of ‘double first-class’ construction: A three-stage global super slacks-based measure analysis,” Education and Information Technologies, vol. 29, no. 10, pp. 12047-12075, 2024, doi: 10.1007/s10639-023-12323-1.

[16] M. Sozen and C. Sozen, “Quality And Efficiency in Higher Education: Assessing The Performance Of Selected Turkish Universities Using The Slack-Based Measure Data Envelopment Analysis,” Journal of Management Theory and Practices Research, vol. 5, no. 2, pp. 277-289. https://dergipark.org.tr/en/pub/jmtpr/article/1595664?issue_id=89248, 2024.

[17] A. Sousa, S. Cardoso P J, and J. Monteiro, “Smart Energy Monitoring for Sustainable Campuses: A Hybrid Anomaly Detection Approach Based on Prophet and Isolation Forest,” Sustainability, vol. 18, no. 5, pp. 2589-2593, 2026, doi: 10.3390/su18052589.

[18] S. Xu, “A method for predicting the result of university employment education transformation based on multi-source information technology,” International Journal of High Speed Electronics and Systems, vol. 34, no. 03, Art. no. 2540024, 2025, doi: 10.1142/S0129156425400245.

[19] Y. Khan, F. Shah S, and M. Ali, “Enhancing imputation accuracy of weighted K-nearest neighbor algorithm using non-Gaussian kernel functions,” International Journal of Data Science and Analytics, vol. 21, no. 1, pp. 72-81, 2026, doi: 10.1007/s41060-025-00910-z.

[20] B. Feng and L. Zhang, “Optimizing the isolation forest algorithm for identifying abnormal behaviors of students in education management big data,” Journal of artificial intelligence and technology, vol. 4, no. 1, pp. 31-39. https://orcid.org/0009-0003-9245-7795, 2024.

[21] H. Xu, G. Pang, Y. Wang, and Y. Wang, “Deep isolated forest for anomaly detection,” IEEE transactions on knowledge and data engineering, vol. 35, no. 12, pp. 12591-12604, 2023, doi: 10.1109/TKDE.2023.3270293.

[22] C. Cao, T. Wei, S. Xu, F. Su, and H. Fang, “Comprehensive evaluation of higher education systems using indicators: PCA and EWM methods,” Humanities and Social Sciences Communications, vol. 10, no. 1, pp. 432-451, 2023, doi: 10.1057/s41599-023-01938-x.

[23] A. Jewsbury P and S. Johnson M, “Principal component analysis on the covariance matrix for data reduction in large-scale assessments,” Large-scale Assessments in Education, vol. 13, no. 1, pp. 30-36, 2025, doi: 10.1186/s40536-025-00264-9.

[24] S. Chen, S. Mohamed Salleh, A. Mohamad Salleh M, and F. Zhou, “Construction of an evaluation index system for teacher education students’ information literacy: an empirical analysis of Chinese postgraduate programs,” Humanities and Social Sciences Communications, vol. 12, no. 1, pp. 1-11, 2025, doi: 10.1057/s41599-025-05561-w.

[25] S. Zhang, T. Zhang, and X. Wang, “AHP-Based Evaluation of discipline-specific information services in academic libraries under digital intelligence,” Information, vol. 16, no. 3, pp. 245-249, 2025, doi: 10.3390/info16030245.

[26] S. Pant, A. Kumar, M. Ram, Y. Klochkov, and K. Sharma H, “Consistency indices in analytic hierarchy process: a review,” Mathematics, vol. 10, no. 8, pp. 1206-1213, 2022, doi: 10.3390/math10081206.

[27] P. Madzik and L. Falat, “State-of-the-art on analytic hierarchy process in the last 40 years: Literature review based on Latent Dirichlet Allocation topic modelling,” PLoS One, vol. 17, no. 5, Art. no. e0268777, 2022, doi: 10.1371/journal.pone.0268777.

[28] Y. Sun, D. Wang, F. Yang, and S. Ang, “Efficiency evaluation of higher education systems in China: A double frontier parallel DEA model,” Computers & industrial engineering, vol. 176, no. 1, pp. 108979-108986, 2023, doi: 10.1016/j.cie.2023.108979.

[29] A. Shero J, S. Al Otaiba, C. Schatschneider, and A. Hart S, “Data envelopment analysis (DEA) in the educational sciences,” The Journal of Experimental Education, vol. 90, no. 4, pp. 1021-1040, 2022, doi: 10.1080/00220973.2021.1906198.

[30] N. Adna, M. Yusoff N S, and S. AZWANI N, “Measuring the performance of Malaysian universities using Charnes, cooper and Rhodes (CCR) and slack-based measure (SBM) models,” Journal of Quality Measurement and Analysis JQMA, vol. 18, no. 1, pp. 1-11. https://www.ukm.my/jqma/wp-content/uploads/2022/03/Paper-1-Norfarziah.pdf, 2022.

[31] A. Emrouznejad, L. Brzezicki, and C. Lu, “Development and evolution of slacks-based measure models in data envelopment analysis: a comprehensive review of the literature,” Journal of Economic Surveys, vol. 39, no. 5, pp. 2106-2133, 2025, doi: 10.1111/joes.12682.

[32] J. Onumanyi A, N. Molokomme D, J. Isaac S, and M. Abu-Mahfouz A, “AutoElbow: An automatic elbow detection method for estimating the number of clusters in a dataset,” Applied Sciences, vol. 12, no. 15, pp. 7515-7526, 2022, doi: 10.3390/app12157515.

[33] R. Hilyawan F, M. Syafrullah, and I. Sudija, “A Systematic Literature Review on the Optimization of K-Means and Agglomerative Clustering for Student Performance Segmentation: A Comparative Analysis of Elbow and Silhouette Methods,” Edcomtech: Jurnal Kajian Teknologi Pendidikan, vol. 10, no. 1, pp. 73-88, 2025, doi: 10.17977/um039v10i12025p73-88.

[34] E. Montano V and F. Schneider S, “Unveiling the Patterns of Employability in Higher Education: A DBSCAN Cluster Analysis of College Programs,” European Journal of Contemporary Education and E-Learning, vol. 2, no. 6, pp. 73-89, 2024, doi: 10.59324/ejceel.2024.2(6).05.

[35] F. Liu, L. Viano, and V. Cevher, “Understanding Deep Neural Function Approximation in Reinforcement Learning via ε-Greedy Exploration,” Advances in Neural Information Processing Systems, vol. 35, no. 1, pp. 5093-5108. https://proceedings.neurips.cc/paper_files/paper/2022/hash/2119b5ac365c30dfac17a840c2755c30-Abstract-Conference.html, 2022.