

ROBUST HOSPITAL EFFICIENCY RANKING: A BAYESIAN AND MACHINE LEARNING INTEGRATED DYNAMIC SLACKS-BASED MEASURE DEA APPROACH

Chaitali Dey*, Jayanth Kumar Thenepalle, Anita Chaturvedi, K J Ghanashyam

Department of Data Analytics and Mathematical Sciences, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bangalore, India

*Correspondence: cpal134@gmail.com

ABSTRACT

Hospitals are complex organizations operating in a dynamic environment, where multiple parameters vary over time. Ranking hospitals by efficiency using traditional methods often overlooks changes in effectiveness and quality. This study proposes a simulation-based Dynamic Slack-Based Measure Data Envelopment Analysis (Dynamic SBM-DEA) framework.

The study is methodological and simulation based, generated using Monte Carlo simulation methods, calibrated from real multispecialty hospital data, healthcare literature, etc. To evaluate intertemporal hospital performance, the methodology incorporates carry-over effects and unwanted outputs and uses bootstrap validation to evaluate the reliability of efficiency estimations. The persistence and transition of performance states across time are examined using Markov chain analysis, and the stability measurements that are obtained are combined with Dynamic SBM efficiency ratings to create a composite performance indicator. Additionally, the framework can capture operational stability, ranking robustness, and predicted efficiency trends inside a single evaluation model thanks to Bayesian forecasting's probabilistic estimations of future performance. The machine learning model attained a testing R^2 of 0.4765 and a cross-validated R^2 of 0.553 ± 0.048 , indicating moderate predictive capability and stable performance. Sensitivity research showed that the composite ranking framework was resilient, with Kendall's tau values ranging from 0.716 to 0.926 across different weighting schemes and Spearman's rank correlations ranging from 0.881 to 0.983.

The findings revealed a set of hospitals with continuously excellent efficiency, stable rankings, and promising future performance. Staffing levels, surgical activity, and clinical results are some of the most important factors influencing efficiency, according to the SHAP study. The development of an integrated Dynamic SBM-DEA, Machine Learning, and Bayesian decision-support framework that offers a transparent, stability-adjusted, and forward-looking benchmarking tool for evidence-based hospital management, resource allocation, and healthcare policy decision-making is the study's primary contribution.

KEYWORDS

DEA, Dynamic efficiency, Hospital ranking, Bayesian forecasting

Managing multispecialty hospitals with limited resources is challenging due to competing demands across departments for staff, equipment, and funding. These constraints often affect service quality, coordination of care, and timely decision-making [1]. The research by [2] examined hospital performance broadly, particularly by considering organizational and management factors alongside quantitative efficiency measures. It provides a perspective on how internal culture, leadership, and staff engagement relate to performance outcomes. However, it is not directly aligned with quantitative efficiency measurement methods, like DEA, or AI-based prediction models. Hospitals face significant challenges in balancing patient demand with available resource capacity [3], made more difficult by the unpredictable nature of demand and inherent resource constraints. Among these, physicians are among the most critical and scarce assets, essential for hospital operations. Beyond ensuring their availability, effective allocation of physicians according to the hospital's case mix is crucial for delivering high-quality care. Collectively, these studies show how multifaceted hospital performance is, but they don't offer much advice on how to combine operational efficiency assessment with analytics for prediction and decision support.

According to the World Health Organization's Global Health Observatory data for 2022, the global average density of medical doctors stands at 17.2 per 10,000 population. Healthcare systems globally face mounting pressure to address escalating demand for medical services. Hospitals grapple with balancing this demand against constrained resources, which include physicians, nurses, medications, medical equipment, instruments, beds, operating rooms, and related infrastructure. A dynamic, nonparametric DEA framework was proposed [4] that uses performance persistence and intertemporal links to jointly quantify the efficiency of an institution over time and identify failure risk. Similarly, a hospital's efficiency should not be assessed based on results from a single year. Hospitals that demonstrate higher average DEA efficiency with lower variability over time, indicated by a smaller coefficient of variation, reflect more stable and sustainable performance. According to the study [5], the performance of public hospitals was evaluated using a three-stage DEA approach, in which initial efficiency scores were adjusted for environmental and managerial factors through regression analysis to obtain more accurate measures of technical efficiency.

The studies by [4, 6] demonstrate the utility of DEA models outside of the healthcare industry. Their research on banks shows how efficiency varies over time and how dynamic DEA can keep an eye on stability and anticipate any institutional collapses. To overcome bias and uncertainty in efficiency estimation, studies by [7, 8] combine DEA with Bayesian techniques. By incorporating past data, Bayesian DEA generates probabilistic efficiency scores and credibility intervals that enhance ranking stability and dependability in performance comparisons. These studies demonstrate how crucial it is to take temporal dynamics and uncertainty into account when evaluating efficiency. But especially when it comes to hospital performance evaluation, they are mainly concerned with efficiency measurement and do not concurrently handle ranking persistence, uncertainty quantification, explainable prediction, and future performance forecasting within a coherent framework.

This work uses Dynamic SBM-DEA with unwanted outputs to measure ranking stability and intertemporal efficiency in multispecialty hospitals. Bootstrap resampling and Bayesian forecasting generate probabilistic efficiency estimates, while machine learning with SHAP finds significant drivers. The suggested paradigm combines dynamic efficiency assessment, ranking persistence analysis, uncertainty modeling, explainable machine learning, and Bayesian forecasting into a unified decision-support framework, in contrast to previous research that looks at these methods separately. The suggested method creates reliable, forward-looking hospital rankings by integrating forecasting, interpretability, uncertainty assessment, and slack-based efficiency analysis. It also fills up a number of methodological holes seen in the current DEA and healthcare benchmarking literature.

Table 1 provides a list of articles based on DEA, Dynamic DEA, and other forms of DEA models used for various purposes, mostly hospital-related features. The literature review table reveals a number of methodological limitations. The majority of current research focuses on measuring efficiency using traditional, dynamic, fuzzy, or network DEA models; however,

there is little integration of efficiency evaluation with future performance forecasts and predictive analytics. Few studies concurrently account for intertemporal efficiency dynamics, ranking stability, and uncertainty quantification, despite the fact that some include quality indicators, patient satisfaction, death rates, or operational aspects. Additionally, explainable machine learning techniques for identifying the primary drivers of hospital efficiency receive little attention in the literature, which mostly focuses on retrospective performance review. Additionally, it is uncommon for current DEA programs to integrate Bayesian forecasting, Markov persistence assessment, and bootstrap-based uncertainty analysis into a single framework. To promote more reliable benchmarking and healthcare decision-making, an integrated methodology that simultaneously assesses efficiency, ranking stability, uncertainty, performance drivers, and future hospital performance is still required.

This study's methodological contribution is the incorporation of well-established analytical methods into a cohesive framework for assessing hospital efficiency. In particular, the suggested framework integrates Dynamic SBM-DEA for efficiency measurement, Markov chain analysis for temporal stability evaluation, Random Forest for predictive validation, Bootstrap resampling for statistical validation, Bayesian updating for information fusion, and SHAP for model interpretability. The main methodological contribution of this study is the methodical integration of these well-established techniques into a unified decision-support framework for hospital efficiency evaluation.

The methodology used to determine the final hospital rankings is presented in Section 2. The analysis's empirical findings are presented in Section 3. Additional analysis and synthesis of these results are given in Section 4. A thorough analysis of the findings in light of the goals of the study and the body of current literature is provided in Section 5. The study's key findings and ramifications are finally summed up in Section 6.

TABLE 1 LITERATURE REVIEW OF HEALTHCARE PERFORMANCE ANALYSIS: INPUTS, OUTPUTS, AND KEY INSIGHTS

Reference	Model Used	Doctors	nurses	beds	operating cost	inpatient	outpatient	surgeries	satisfaction	Mortality rate	Findings
[9]	Dynamic Network DEA	√	√	√	-	-	√	-	-	-	Demonstrated that hospital efficiency is better evaluated by looking at internal sub-processes (clinical vs. administrative) rather than a "black box" approach.
[10]	Dynamic Network DEA	-	-	√	-	-	√	-	√	-	Established a longitudinal link between efficiency and quality, suggesting that efficient hospitals often maintain higher quality standards over time.
[11]	DEA with Quality Variables	√	√	√	-	-	-	√	-	-	Found that improvements in operational efficiency generally led to improved quality, though the relationship varies by hospital size.
[12]	CCR DEA	√	-	-	√	-	-	-	-	√	Analyzed regional healthcare in Slovakia, finding that regions with higher spending did not always correlate with better population health outcomes.
[13]	DEA with Environmental Variables	√	√	√	-	√	√	-	-	-	Factors such as the ratio of specialized technicians and the proportion of outpatient visits significantly influenced these efficiency levels.
[14]	Network DEA	√	√	√	√	-	-	-	-	√	quantifies the efficiency of health systems across 37 OECD countries by analyzing internal sub-processes
[15]	Window-DEA	√	√	√	√	-	-	√	-	-	Conducted a longitudinal (Window-DEA) evaluation of the healthcare sector's efficiency in Poland.
[16]	Credibility-based fuzzy window DEA.	√	√	√	√	√	√	-	-	-	Used credibility-based fuzzy window DEA for dynamic performance assessment of hospitals.
[17]	Systematic review of features used	√	√	√	√	√	√	-	-	√	A systematic review that identified common trends and methodological gaps in health system efficiency literature.
[18]	Two-stage DEA approach	√	√	√	√	-	-	-	-	√	Evaluated the efficiency of healthcare systems in Arab countries using a two-stage DEA approach.

[19]	DEA methodology tot.	-	-	-	-	√	√	√	√	-	Provided an empirical analysis of primary healthcare efficiency within a specific district using standard DEA.
[20]	Network-DEA model.	√	√	-	√	-	-	-	√	-	Demonstrated how the quality of care and patient access directly impact overall hospital efficiency.
[21]	Dynamic Network DEA	-	-	√	√	√	√	√	-	-	Applied a Dynamic Network DEA approach to evaluate hospital performance over multiple periods.

METHODOLOGY

DATA AND SAMPLE DESCRIPTION

This study considers multispecialty hospitals in India as the decision-making units for ranking. Hospital-level panel data covering inputs, outputs, undesirable outcomes, and carry-over variables over multiple years are very limitedly available in a consistent and complete form, especially for multispecialty hospitals. To overcome this limitation while preserving realism and methodological rigor, this study employs a hierarchical Monte Carlo simulation approach to generate a synthetic but policy-relevant hospital panel dataset [22]. To capture changes in performance and intertemporal links in hospital operations, the analysis spans an eight-year study period. Operational benchmarks published in Indian healthcare literature, NABH accreditation requirements, National Health Profile reports, and annual reports of top multispecialty hospitals were used to calibrate the synthetic dataset. While service utilization metrics such as outpatient visits, inpatient admissions, surgical volumes, and mortality rates were obtained from published data for tertiary-care hospitals, staffing levels were determined using average doctor-to-bed ratios of roughly 1:8 and nurse-to-bed ratios of 1.5-2.0. The Monte Carlo simulation's probability distributions and parameter ranges were established using these empirical benchmarks, as shown in Table 3. This ensured that the database was appropriate for Dynamic SBM-DEA and the ensuing ranking stability analysis.

DYNAMIC SBM-DEA EFFICIENCY ESTIMATION

Using a Slack-based DEA framework, the investigation begins by calculating annual efficiency scores. Four input factors (Doctors, Nurses, Beds_Input, Operating_Expenditure_Cr) and four output variables (IPD_Admissions, OutpatientsOPD_Visits, Surgeries, Satisfaction_Pct) are used to assess hospitals annually, as shown in Table 2. To maintain temporal variety, efficiency is computed independently for every year.

TABLE 2 DYNAMIC DEA VARIABLES

Type	Components
Inputs	Doctors, Nurses, Beds_Input, Operating_Expenditure_Cr
Outputs	IPD_Admissions, OutpatientsOPD_Visits, Surgeries, Satisfaction_Pct
Carryovers	Beds_Carryover, Technology_Index, Finance_Strength_Index
Undesirables	Mortality_Rate_Pct

The hospital's physical infrastructure and service capacity that will be accessible in later years are represented by Beds_Carryover. The original dynamic SBM framework of Tone K. Tsutsui and Miki Tsutsui, where carry-overs represent resources transferred between periods. This is the reason to keep beds in Input as well as Carryovers. Beds_Carryover, Technology_Index, and Finance_Strength_Index were chosen as desirable carry-over factors because of their long-lasting impact on hospital performance over a number of time periods. These resources endure and continue to impact future operating capacity and efficiency, in contrast to traditional inputs that are depleted within a year.

The Technology Index and Finance Strength Index were incorporated as carry-over variables because their effects extend beyond a single period and influence future hospital performance. The level of medical equipment, digital infrastructure, and technical capabilities that improve patient outcomes, service quality, and operational efficiency over time is represented by the Technology Index. In a similar vein, the hospital's retained financial capacity and investment potential,

which enable future growth, technology adoption, and service enhancements, are reflected in the Finance Strength Index. These strategic resources are appropriate carry-over variables in the Dynamic SBM-DEA framework because they create an intertemporal connection between previous managerial choices and future efficiency outcomes as they are gathered and transferred over time. Mortality Rate was viewed as an undesirable output since it represents patient safety and clinical results; lower levels signify greater operational performance and healthcare quality. A more thorough evaluation of long-term hospital sustainability and efficiency is made possible by taking these factors into account.

Dynamic DEA extends static DEA to evaluate hospital efficiency across multiple periods T , incorporating carry-over criteria such as the number of beds, thereby enabling stable ranking via period-wise and overall scores. Consider n hospitals (DMUs, $j = 1, 2, \dots, n$), over T periods, with m inputs $x_{ijt} > 0$, for example, staff, beds at period t , s outputs $y_{rjt} > 0$, for example, inpatients, outpatients, revenue, p carry-overs $z_{kjt} > 0$, such as capital stock, beds, technology, etc.

Let $\lambda_{jt} > 0$ be the intensity (peer weight) of DMU.

The efficiency formulation is appropriate for a dynamic SBM model within the DEA framework and is well-suited to dynamic healthcare efficiency analysis. By removing the evaluated DMU from the reference set, the model minimizes average inefficiency over time and permits efficiency scores to exceed 1, enabling discrimination among efficient units [23].

Slack variables are input slack ($s_{it}^- \geq 0$), output slack ($s_{rt}^+ \geq 0$), carry-over slacks ($s_{kt}^{c-}, s_{kt}^{c+} \geq 0$). The period-wise constraints

(a) for each input i and period t

$$x_{iot} = \sum_{j=1}^n \lambda_{jt} x_{ijt} + s_{it}^- \quad (1)$$

(b) for each output r and period t :

$$y_{rot} \geq \sum_{j=1}^n \lambda_{jt} y_{rjt} - s_{rt}^+ \quad (2)$$

(c) for each carry over (dynamic link) constraints for variable k and period t :

$$y_{rot} \geq \sum_{j=1}^n \lambda_{jt} y_{rjt} - s_{kt}^{c+} - s_{kt}^{c-} \quad (3)$$

The overall dynamic SBM efficiency score for DMU o is obtained by solving

$$\min_{\lambda, s} \rho_o$$

$$\text{where } \rho^{SE} = \frac{1}{T} \sum_{t=1}^T \left[\frac{1 - \frac{1}{m} \sum_{i=1}^m \frac{s_{it}^-}{x_{iot}}}{1 + \frac{1}{s} \sum_{r=1}^s \frac{s_{kt}^+}{y_{kot}}} \right] \quad (4)$$

This objective $\min \rho^{SE}$ computes ρ^{SE} as a period-averaged SBM score.

Dynamic hospital ranking and stability

Year-wise ranking

$$R_{jt} = \text{rank descending of } \rho_{jt}^{SE} \quad (5)$$

Average rank

$$\bar{R}_j = \frac{1}{T} \sum_t R_{jt} \quad (6)$$

Rank stability (coefficient of variation)

$$CV_j = \frac{\sigma(R_{jt})}{\bar{R}_j} \quad (7)$$

Lower CV implies a more stable ranking.

MACHINE LEARNING MODELS

The study highlights that machine learning models analyze large volumes of historical and real-time data to identify patterns and predict future needs. These predictive capabilities allow firms to plan production, inventory, and resource allocation more accurately. As a result, companies can reduce overproduction, minimize excess inventory, and lower operational costs [24].

Bootstrap uncertainty in efficiency and ranking

This approach enables statistical inference for Dynamis SBM-DEA results, indicating whether observed differences in efficiency and ranking are robust or driven by sampling variation. Consequently, hospitals whose confidence intervals overlap substantially or whose rank distributions are highly dispersed are interpreted as having less certain performance positions. In contrast, narrow intervals and stable ranks indicate reliable efficiency assessments.

For each period t

Resample hospitals with replacement: dataset D^*

Re-estimate SE-DSBM: ρ_{jt}^{SE*} . Repeat the following steps B times.

$$\text{Bootstrap mean efficiency: } \widehat{\rho_{jt}^{boot}} = \frac{1}{B} \sum_{b=1}^B \rho_{jt}^{*(b)} \quad (8)$$

$$\text{Bias-corrected efficiency: } \rho_{jt}^{BC} = \rho_{jt}^{SE} - (\widehat{\rho_{jt}^{boot}} - \rho_{jt}^{SE}) \quad (9)$$

$$\text{Confidence interval: } CI_{jt} = [\rho_{jt}^{*(\alpha/2)}, \rho_{jt}^{*(1-\alpha/2)}] \quad (10)$$

Rank uncertainty: From each bootstrap sample, $R_{jt}^{*(b)}$, rank variance can be obtained

$$\text{var}(R_{jt}) = \frac{1}{B} \sum (R_{jt}^{*(b)} - \overline{R_{jt}})^2 \quad (11)$$

Markov rank transition analysis

Markov rank transition analysis uses dynamic efficiency rankings to monitor how hospitals move across performance levels over time. Transitions create a probability matrix, and hospitals are categorized into rank states (such as quartiles). Off-diagonals indicate mobility, while diagonal values indicate persistence. Mobility indices and steady-state distributions show whether rankings are dynamic, convergent, or stable. To examine the temporal dynamics of hospital performance, efficiency scores from the Dynamic super-efficiency SBM model ρ_{jt}^{SE} were converted into annual ranks R_{jt} for each hospital $j = 1, \dots, n$ and year $t = 1, \dots, T$.

Let E_j = Dynamic SBM Efficiency score of hospital j

For each hospital j , remove its own efficiency score from the reference set and determine the maximum efficiency among the remaining hospitals:

$$E_{\max(-j)} = \max\{E_i: j \neq i\}$$

The Super-Efficiency score is then computed as:

$$\rho_{jt}^{SE} = \frac{E_i}{E_{\max(-j)}} \quad (12)$$

where:

ρ_{jt}^{SE} = Super-Efficiency score of hospital j

E_j = Efficiency score of hospital j

$E_{\max(-j)}$ = Maximum efficiency score among all hospitals excluding hospital j

Ranks were then grouped into K ordered performance states, for example, quartiles, using a mapping $s_{jt} = g(R_{jt})$, where $s_{jt} \in \{1, \dots, K\}$. One-period transitions between consecutive years were counted as $N_{ab} = \{j, t: s_{jt} = a, s_{j,t+1} = b\}$, and the Markov transition probability matrix was obtained by row normalization, $P_{ab} = N_{ab} / \sum_{b=1}^K N_{ab}$, with $\sum_b P_{ab} = 1$. Diagonal

elements P_{aa} measure rank persistence, while off-diagonal elements capture upward and downward mobility; the long-run distribution of hospitals across rank states was derived from the stationary vector π satisfying $\pi = \pi P$ and $\sum_a \pi_a = 1$.

Composite Ranking

In the present study, with 100 hospitals evaluated across four inputs, four desirable outputs, one undesirable output, and three carry-over link variables, this degeneracy risk is non-trivial. To address it, the study presents a composite scoring framework that adds two orthogonal dimensions, temporal behavioural consistency and statistical rank confidence, to the Dynamic SBM efficiency estimate. These dimensions are theoretically independent of the DEA frontier geometry. The composite score for the hospital i was calculated as:

$$CS_i = 0.5 E_i + 0.25 M_i + 0.25 B_i \quad (13)$$

where the weight of 0.50 assigned to E_i reflects the primacy of technical efficiency as the core DEA construct, while the equal weights of 0.25 on M_i and B_i encode the assumption that temporal consistency and statistical confidence contribute symmetrically to the reliability of a hospital's ranking.

ML validation of DEA and determinants (Random Forest + SHAP)

To validate the structural consistency of the dynamic super-efficiency DEA results and identify key determinants of sustainable hospital performance, a supervised machine learning model was estimated using DEA inputs, desirable outputs, undesirable outputs, and carry-over variables as predictors. In order to gain a deeper understanding of the factors that influence efficiency, many studies combined machine learning with DEA. After efficiency scores are calculated using DEA, ML models, such as RF, are used to forecast efficiency levels and pinpoint the primary operational factors affecting performance. This study also makes use of AI tools like SHAP [25, 26].

The dataset was divided into 640 training samples (80%) and 160 testing samples (20%). To assess model stability and lower the possibility of overfitting, a 10-fold cross-validation technique was used. The coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were used to evaluate the model's performance.

For each hospital j and year t , the feature vector was defined as

$$\mathbf{X}_{jt} = [x_{ijt}, y_{rjt}, b_{pjt}, z_{kjt}], \quad (14)$$

with the DEA super-efficiency score as the target $Y_{jt} = \rho_{jt}^{SE}$.

A Random Forest regression model $\hat{\rho}_{jt}^{RF} = f_{RF}(\mathbf{X}_{jt})$ was trained by minimizing prediction error $\min \sum_{j,t} (\rho_{jt}^{SE} - \hat{\rho}_{jt}^{RF})^2$; agreement between the DEA and ML estimates was assessed via correlation $\text{Corr}(\rho^{SE}, \hat{\rho}^{RF})$ and error metrics. Determinants of efficiency were then interpreted using SHAP values, where the contribution of the variable k to a prediction is

$$\phi_k = \sum_{S \subseteq F \setminus \{k\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f(S \cup \{k\}) - f(S)] \quad (15)$$

and overall importance was summarized by the mean absolute SHAP $I_k = \frac{1}{N} \sum |\phi_k|$, identifying variables most influential for sustained hospital rankings.

BAYESIAN FUSION MECHANISM

The study begins with Dynamic SBM DEA to estimate intertemporal hospital efficiency and slack structures. In subsequent stages, including bootstrap and Bayesian forecasting, the same Dynamic SBM efficiency scores and their uncertainty distributions are propagated forward. To quantify intertemporal hospital efficiency and slack structures, the investigation starts with Dynamic SBM-DEA. The Dynamic SBM efficiency scores and their corresponding uncertainty distributions are carried over into further phases, such as machine learning prediction and bootstrap analysis. Robust future efficiency estimates are then produced by combining predicted data from the Random Forest model with historical efficiency evidence using a Bayesian fusion framework.

Prior Distribution

The DEA bootstrap efficiency estimates serve as the prior distribution for each hospital. Let θ_h denote the latent efficiency of hospital h . The prior is specified as the DEA efficiency defines the prior

$$\theta_h^{prior} \sim \mathcal{N}(\mu_h^{DEA}, \sigma_p^2) \quad (16)$$

where,

μ_h^{DEA} = mean bootstrap SBM efficiency, $\sigma_p = 0.004$,

bounded to feasible DEA range: $\theta_h^{prior} \in [0.95, 1]$

In addition to providing a statistically supported prior distribution, the bootstrap-derived variance captures the uncertainty related to efficiency estimation.

Likelihood Specification

Random Forest predictions are treated as noisy observations of future hospital efficiency. The likelihood function is specified as

$$\theta_h^{like} \sim \mathcal{N}(\theta_h^{RF}, \sigma_L^2) \quad (17)$$

where θ_h^{RF} is the Random Forest predicted efficiency score and σ_L^2 denotes the prediction error variance estimated from out-of-sample testing performance. The likelihood therefore incorporates information from operational, financial, infrastructure, and clinical variables used in the machine learning model.

Hyperparameter Selection

The hyperparameters σ_p and σ_L ascertain the relative impact of Random Forest and DEA data throughout the fusion process. The bootstrap confidence intervals of DEA efficiency estimates are used to calculate the prior standard deviation (σ_p), whereas the likelihood standard deviation (σ_L) is calculated using the residuals of the Random Forest forecast. The weighting system was chosen to incorporate predictive data while highlighting past efficiency performance. The relative contributions of the Dynamic SBM-DEA prior and the Random Forest likelihood were systematically varied across five weight combinations (0.5,0.5), (0.6,0.4), (0.7,0.3), (0.8,0.2), and (0.9,0.1) to conduct a sensitivity analysis. All weight combinations produced exceptionally high Spearman rank correlation coefficients ($\rho = 0.998099$ to 1.000000 , $p < 0.001$), as Table 3 illustrates. This suggests that the suggested Bayesian fusion framework is very robust and that the final hospital rankings are essentially unaffected by moderate changes in the weight allocation. This indicates that the model is stable and independent of specific weight selections. The 0.6 (DEA) and 0.4 (Random Forest) weighting was chosen as the baseline because it produced the best agreement with the reference ranking ($\rho = 1.000$) while preserving a suitable balance between the two information sources. The weighted combination of DEA prior and RF likelihood is

$$\theta_h^{post} = 0.6\theta_h^{prior} + 0.4\theta_h^{like} \quad (18)$$

where the weights 0.6 and 0.4 reflect greater confidence in observed DEA performance while allowing machine learning forecasts to influence future efficiency estimation. The ML component (0.4) acts as a corrective signal from the demand side, while the DEA component (0.6) is given more weight because it directly reflects operational efficiency from realized performance.

TABLE 3 SENSITIVITY EVALUATION OF BAYESIAN FUSION WEIGHTS

Weight pair	Equation	Spearman(ρ)	p-value	Stability
(0.5,0.5)	$\theta_h^{post} = 0.5\theta_h^{prior} + 0.5\theta_h^{like}$	0.999828	0	Very High
(0.6,0.4)	$\theta_h^{post} = 0.6\theta_h^{prior} + 0.4\theta_h^{like}$	1	0	Baseline
(0.7,0.3)	$\theta_h^{post} = 0.7\theta_h^{prior} + 0.3\theta_h^{like}$	0.999801	0	Very High
(0.8,0.2)	$\theta_h^{post} = 0.8\theta_h^{prior} + 0.2\theta_h^{like}$	0.999177	0	Very High
(0.9,0.1)	$\theta_h^{post} = 0.9\theta_h^{prior} + 0.1\theta_h^{like}$	0.998099	0	Very High

Posterior Estimation Procedure

The Random Forest likelihood and the DEA-based prior are combined to get the posterior efficiency estimate. After taking into account both past performance and predictive analytics, the resulting posterior distribution shows the updated belief of future hospital efficiency. Here, the efficiency is restricted to the range:

$$\theta_h^{future} = \min(1, \max(0.7, \theta_h^{post})) \quad (19)$$

Thus, the final forecast (mean over samples):

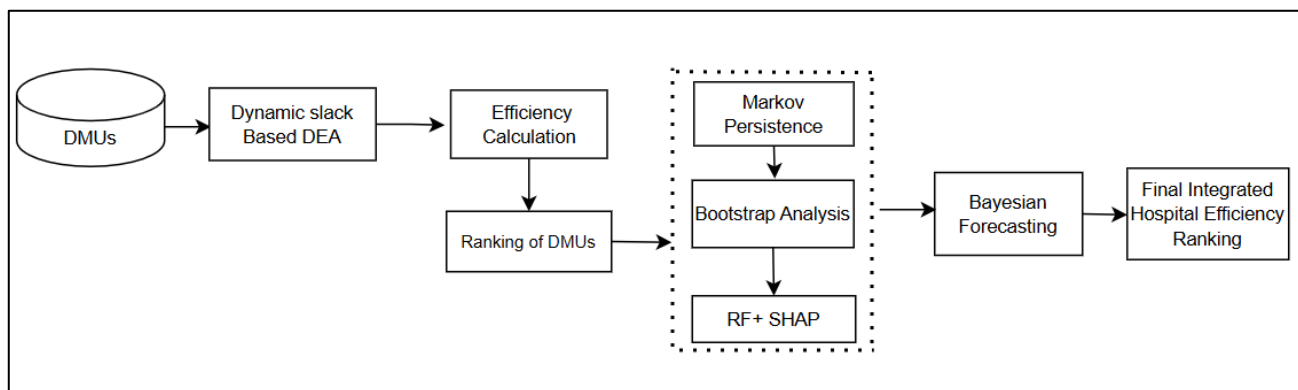
$$\hat{\theta}_h^{future} = E[\theta_h^{future}] \quad (20)$$

Hospitals will be ranked in descending order $\hat{\theta}_h^{future}$.

Uncertainty Quantification

Monte Carlo sampling is used to create posterior distributions from the prior and likelihood distributions in order to assess uncertainty in future efficiency estimates. Posterior means, standard deviations, and 95% credible intervals are computed for every hospital. Wider credible intervals imply more uncertainty in future rankings, while narrow credible intervals show steady and trustworthy efficiency forecasts. This uncertainty assessment offers a probabilistic interpretation of future hospital performance and supplements the bootstrap confidence intervals derived from Dynamic SBM-DEA.

FIGURE 1. INTEGRATED WORKFLOW



RESULTS

A hierarchical Monte Carlo simulation framework was used to create a synthetic panel dataset with 100 hospitals observed every year from 2016 to 2023 (800 hospital-year observations). Baseline characteristics, such as size category, bed capacity, technology index, and financial carryover, were initially sampled for each hospital from empirically calibrated distributions that matched multispecialty hospital standards approved by NABH. Then, based on these baseline characteristics, annual staffing, utilization, spending, satisfaction, and mortality factors were simulated.

Descriptive statistics, distributional analysis, and empirical benchmark (Table 4) comparison were used to evaluate the synthetic dataset's validity. The produced observations showed central tendencies, variability, and realistic ranges that were in line with operational traits documented for multispecialty hospitals in India. The simulated dataset's face validity and practical realism were supported by mean values and dispersion measurements that matched NABH-based staffing norms, hospital use statistics, and published healthcare performance reports. Figure 2 shows the histograms for each of the normal and uniform variables in order to have a better understanding of each variable's distribution and to verify the assumptions stated for the Kolmogorov-Smirnov test visually.

TABLE 4. MONTE CARLO SIMULATION MODEL PARAMETERIZATION USING EMPIRICAL BENCHMARKS AND VALIDATION

Variable	Initial Simulation Distribution	Literature range	Simulated Mean	Simulated Range	Empirical Source	Validation
Doctors per bed	Normal	0.3-0.5	0.43	0.4-0.59	NABH staffing guidelines; annual reports of multispecialty hospitals	Consistent with staffing levels in large multispecialty hospitals Reflects realistic nursing workforce distribution
Nurses per bed	Normal	0.8-1.5	0.99	0.80-1.44	NABH staffing guidelines	Consistent with admission volumes reported by medium-to-large multispecialty hospitals Reflects realistic outpatient service demand and utilization patterns
Operating Expenditure	Uniform	Rs 40- Rs 200 Cr	Rs. 111.5	Rs41.4- Rs202.2	Hospital annual reports; Indian healthcare cost studies	Consistent with expected expenditure levels
Inpatients (IPD Admissions)	Normal	10000-50000	30178	11142-48193	Indian hospital utilization studies; National Health Profile	Consistent with accreditation hospital performance
Outpatients (OPD Visits)	Normal	1000-1800	1303	677-2381	National Health Profile; hospital annual reports	Within reported clinical outcome ranges
Patient Satisfaction	Uniform	70-95%	90.80%	71-98%	Published patient satisfaction studies in Indian hospitals	Reflects progressive technology adoption
Mortality Rate	Uniform	1-4%	2	1-4%	Tertiary-care hospital outcome studies	Consistent with occupancy levels observed in efficiently managed multispecialty hospitals
Technology Index	Uniform	50-100	80.65	51-100	NABH infrastructure scoring; calibrated to Capital_Stock_Cr	
Bed Occupancy Rate	Uniform	60-85%	69.38%	37.6-121.6%	WHO & MoHFW target; upper tail >100% reflects day-care turnover	

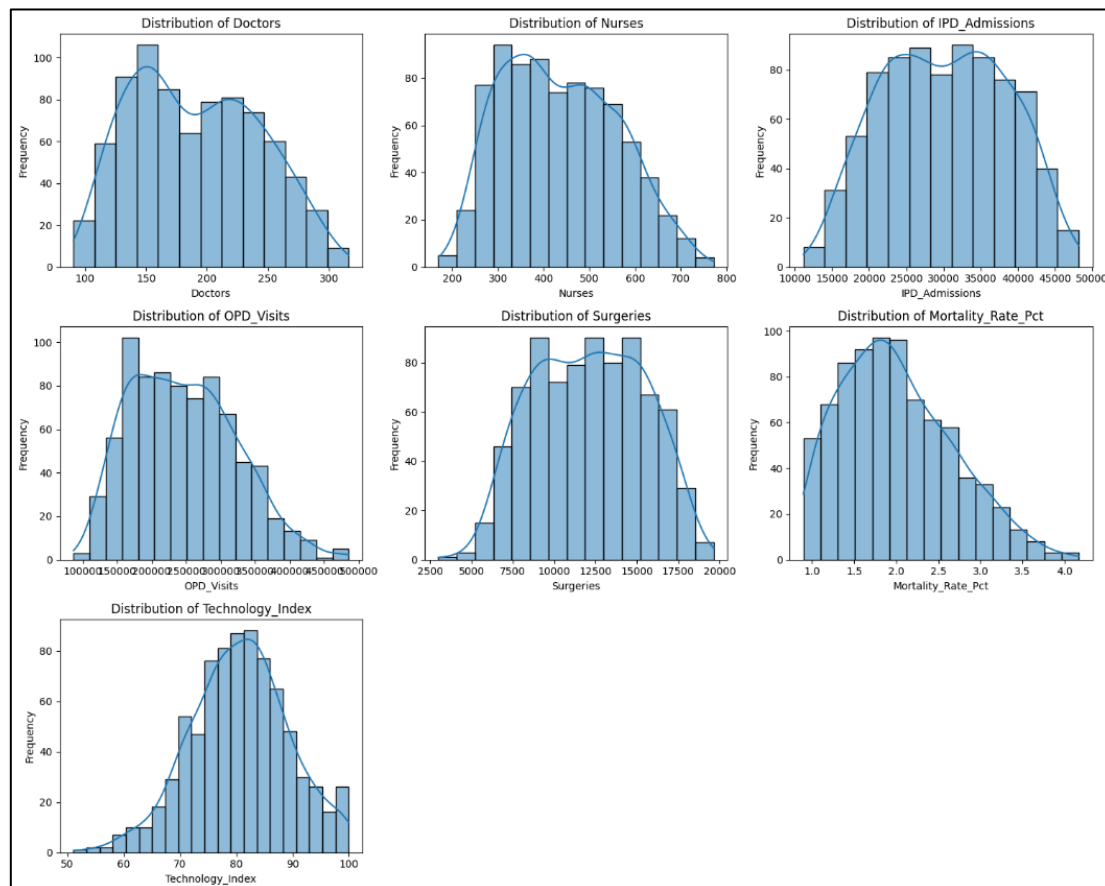
MoHFW: Ministry of Health and Family Welfare, Government of India.

NABH: National Accreditation Board for Hospitals & Healthcare Providers

WHO: World Health Organization

Empirical benchmarks from NABH standards, National Health Profile reports, hospital annual reports, and published healthcare studies were used to calibrate and validate the synthetic dataset. The dataset offers a realistic setting for assessing the suggested Dynamic SBM-DEA, machine learning, and Bayesian forecasting framework, according to the congruence between the generated values and reported operational features of Indian multispecialty hospitals, as described in Table 3.

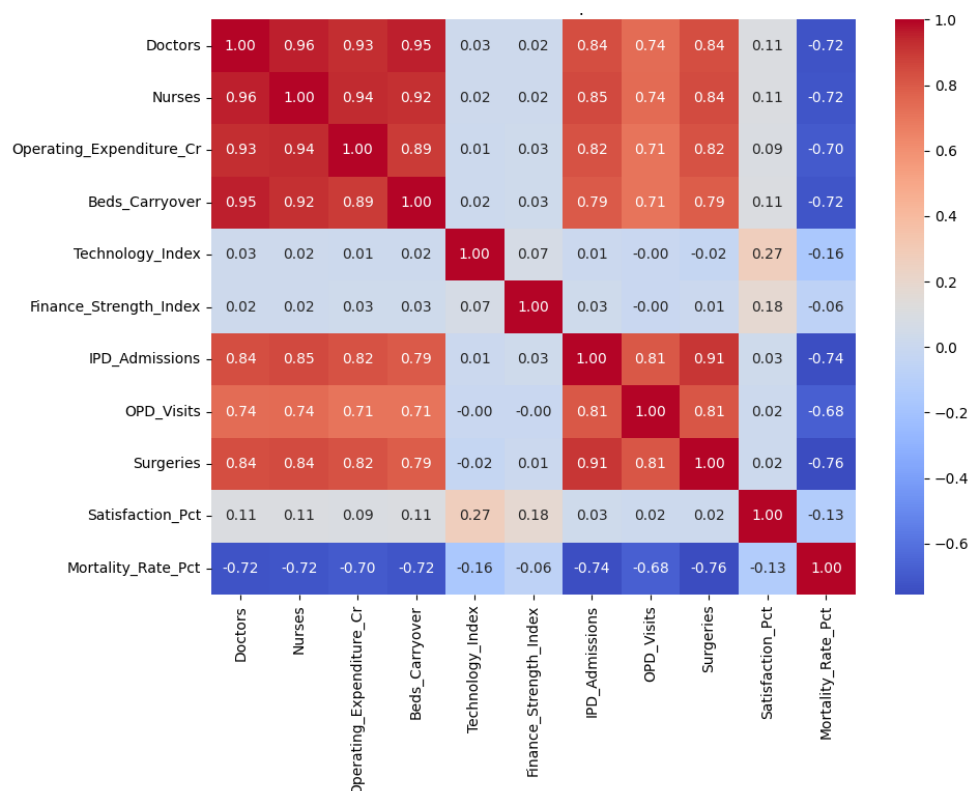
FIGURE 2: VISUALIZING VARIABLE DISTRIBUTIONS



The results of the correlation study in Figure 3 are clinically consistent and validate the DDEA model's structure. The input variables have the largest positive correlations, with Doctors exhibiting substantial connections with Nurses ($r = 0.96$), Beds_Carryover ($r = 0.95$), and Operating_Expenditure_Cr ($r = 0.93$), suggesting bed capacity as the main structural driver of hospital input volume. The volume-outcome hypothesis is supported by the output variables' high internal clustering, with IPD_Admissions, Surgeries, and OPD_Visits intercorrelated at $r = 0.81$ – 0.91 . Clinical consistency is further confirmed by significant negative correlations between Mortality_Rate_pct and Surgeries ($r = -0.76$), IPD_Admissions ($r = -0.74$), and Doctors ($r = -0.72$), suggesting that higher-capacity hospitals achieve better patient safety results.

Variables were restricted to important operational, clinical, infrastructure, and financial aspects that were backed by NABH standards and healthcare literature in order to allay dimensionality concerns. Technology, finance, satisfaction, and mortality showed lesser correlations, suggesting different performance aspects, but staffing, bed capacity, expenditure, and service volumes showed high positive connections demonstrating hospital scale effects. These connections validate the dataset's eligibility for Dynamic DEA analysis.

FIGURE 3: CORRELATION MATRIX OF HOSPITAL VARIABLES



INTERPRETATION OF KOLMOGOROV-SMIRNOV (KS) GOODNESS-OF-FIT RESULTS

The consistency of the simulated variables with their assumed normal distributions was assessed using the Kolmogorov-Smirnov goodness-of-fit test. The related KS statistics were rather small (0.050–0.078), as shown in Table 5, indicating little deviation from normalcy, even though all variables produced statistically significant p-values ($p < 0.05$). The KS test is extremely sensitive to small distributional changes because of the study's huge sample size. Additionally, a visual analysis of the histograms (Fig. 2) verified that the derived variables closely match the anticipated patterns documented in hospital performance studies and Indian healthcare literature. Consequently, the findings imply that the synthetic dataset maintains genuine variability and heterogeneity among hospitals while fairly approximating the anticipated distributions.

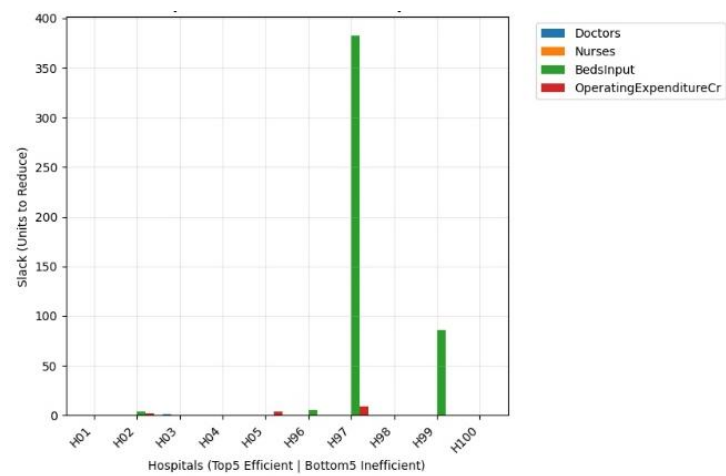
TABLE 5 KOLMOGOROV-SMIRNOV (KS) GOODNESS-OF-FIT RESULTS

Variable	KS Statistic	P-value
Doctors	0.0775	0.0001
Nurses	0.0733	0.0003
IPD Admissions	0.0501	0.035
OPD Visits	0.0586	0.0078
Surgeries	0.0544	0.0168

The synthetic dataset was calibrated using operational benchmarks given in Indian healthcare literature, NABH certification requirements, National Health Profile statistics, and yearly reports of multispecialty hospitals, even though the study is mainly methodological and simulation based. The goal was to create realistic hospital scenarios that maintain the structural relationships, variability, and operational characteristics seen in practice rather than to imitate the performance of particular hospitals. As a result, the synthetic data offer a reliable testing environment for assessing the suggested Dynamic SBM-DEA, bootstrap, Markov transition, machine learning, and Bayesian forecasting framework.

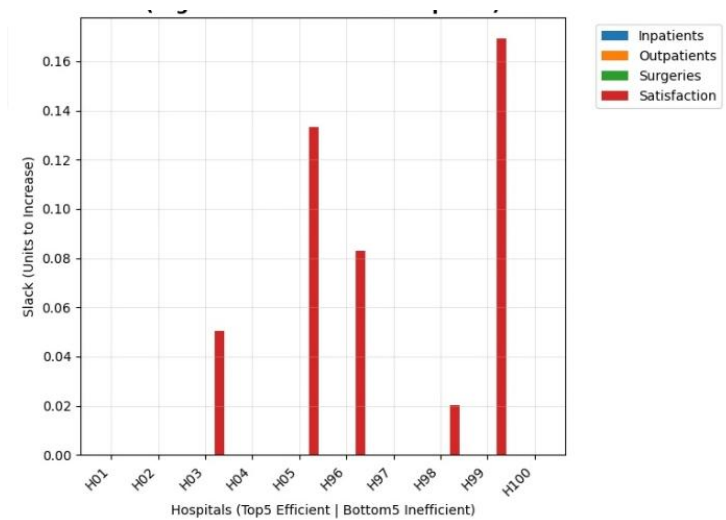
Slacks-Based Measure (SBM) DEA models, introduced by Tone (2002), provide a non-radial efficiency measure that directly minimizes input excesses and output shortfalls relative to the frontier. Unlike radial CCR/BCC models that proportionally adjust inputs or outputs and may leave residual slacks, SBM evaluates inefficiency explicitly. The measure is unit-invariant, monotone in inefficiency, and defined by a DMU's reference set, which supports robust application in network and dynamic extensions such as the Tone–Tsutsui framework.

FIGURE 4: INPUT SLACKS- REDUCTION TARGETS



Based on SBM-DEA, Fig. 4 shows input slack levels that signify excess resource consumption across the top 5 efficient and bottom 5 inefficient hospitals. Across all input categories (Doctors, Nurses, Beds_Input, and Operating_Expenditure_Cr), efficient hospitals (leftmost 5 bars) exhibit almost zero slack values, indicating optimal resource allocation with little waste relative to the efficiency frontier.

FIGURE 5: OUTPUT SLACKS – EXPANSION TARGETS



Output slack levels, which indicate unrealized production potential across hospital groups, are shown in Fig. 5. Near-maximum production in inpatient, outpatient, surgical, and satisfaction results is demonstrated by efficient hospitals, which exhibit very little slack. While greater inpatient and surgical gaps indicate lost revenue potential, moderate outpatient slack shows sufficient use. Maintaining a suitable ratio between the total number of model variables and the number of decision-making units (DMUs) is essential to the discriminatory power of DEA. The current study makes use of a panel dataset that includes 800 hospital-year observations from 100 multispecialty hospitals across eight years. Four inputs, four desired outputs, three carry-over variables, and one unwanted output make up the Dynamic SBM-DEA model. As a result, the number of observations significantly surpasses the model dimensionality, meeting widely recognized DEA sample-size requirements and lowering the possibility that an excessive percentage of hospitals will be categorized as efficient.

TABLE 6 HOSPITAL RANKING ACROSS YEARS

Year	2016	2017	2018	2019	2020	2021	2022	2023
Hospital_ID								
H009	21	1	10	37	5	11	1	1
H013	23	15	15	12	6	22	38	29
H016	9	2	5	25	10	6	2	2
H021	18	13	30	9	18	16	10	9
H022	16	14	6	18	4	3	9	5
H037	6	8	44	51	7	10	5	65
H039	1	47	1	4	1	2	50	12
H040	28	19	27	2	24	13	11	10
H048	44	3	25	28	14	25	3	3
H050	30	7	21	19	8	12	6	17
H052	14	10	17	8	3	17	33	15
H057	5	9	20	15	13	9	16	11
H062	4	21	22	6	20	15	8	6
H063	7	4	8	14	2	57	7	13
H070	20	49	2	33	23	65	53	16
H072	3	17	11	11	11	21	4	19
H077	15	59	90	3	86	1	79	28
H081	2	5	3	10	41	43	20	4
H084	10	50	12	5	74	4	31	26
H091	8	18	33	7	16	8	15	20

Table 6 explains the variation in hospital rankings over the years. Based on this information, Dynamic SBM efficiency and super-efficiency scores are calculated. The top 20 Hospitals are shown above.

The Dynamic SBM-DEA efficiency scores, super-efficiency scores, and matching hospital rankings based on performance across various time periods are displayed in Table 7 below.

TABLE 7 DYNAMIC SBM DEA RANKING AND SUPER EFFICIENCY

Hospital_ID	Dynamic_SBM_Efficiency	Super_Efficiency	Super_Rank
H039	0.756107	1.061449	1
H009	0.712335	0.942108	2
H016	0.706908	0.934931	3
H022	0.64497	0.853014	4
H063	0.626936	0.829162	5
H081	0.626666	0.828806	6
H062	0.622437	0.823212	7
H072	0.611881	0.809252	8
H057	0.594665	0.786482	9
H040	0.590096	0.78044	10
H052	0.589282	0.779363	11
H021	0.584846	0.773496	12
H091	0.583404	0.771589	13
H048	0.581759	0.769413	14
H050	0.580761	0.768093	15
H037	0.563658	0.745473	16
H084	0.555012	0.734038	17
H013	0.541276	0.715872	18

H077	0.534368	0.706736	19
H070	0.522487	0.691023	20

Table 8 describes the descriptive statistics for Dynamic SBM. The mean efficiency is 0.8337, with a standard deviation of 0.1065, indicating that the Dynamic SBM efficiency scores show moderate variation in hospital performance across the sample. Instead of grouping all hospitals close to the efficiency frontier, the efficiency values ranged from 0.6461 to 1.0000, indicating that the Dynamic SBM-DEA model has sufficient discriminatory ability to differentiate between reasonably efficient and less efficient hospitals.

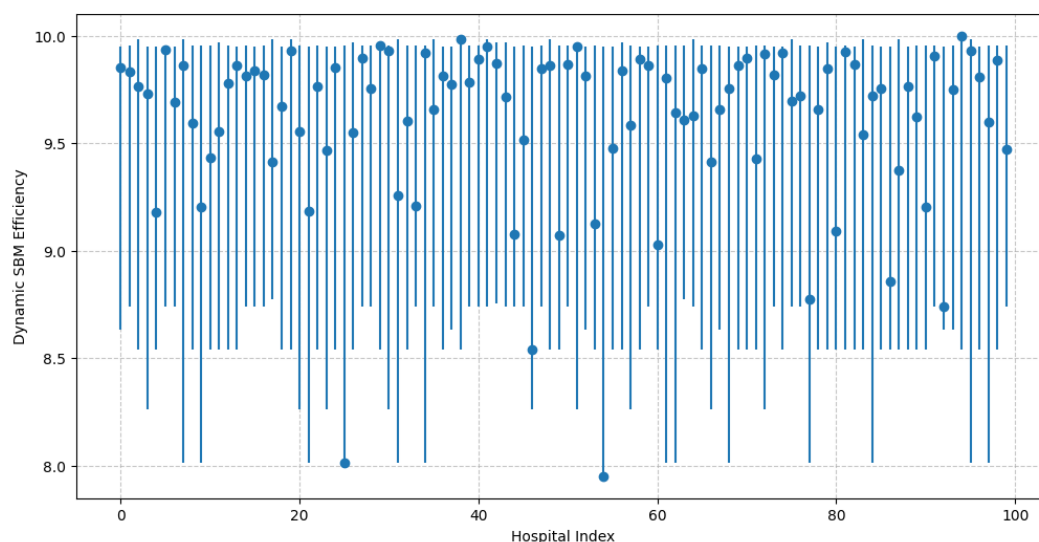
TABLE 8 DESCRIPTIVE STATISTICS FOR DYNAMIC SBM EFFICIENCY

Statistic	Value
Count	100
Mean	0.8337
Standard Deviation	0.1065
Minimum	0.6461
25% (Q1)	0.7372
50% (Median)	0.8265
75% (Q3)	0.9293
Maximum	1

According to the Dynamic SBM-DEA scores, there was moderate fluctuation across hospitals, indicating notable variations in operational effectiveness.

The statistical reliability of the Dynamic SBM-DEA efficiency ratings and ranks was evaluated using bootstrap resampling. Each hospital's rank distributions and confidence intervals were produced by the bootstrap process, allowing for the assessment of whether observed ranking disparities were significant or due to sampling variation. Bootstrap analysis was very helpful in illustrating the uncertainty surrounding hospital rankings and in identifying ranks that were statistically not distinguishable from efficiency ratings. With the mean Bootstrap_SD of 0.045, the bootstrap analysis (B = 500) produced an average 95% confidence interval width of 0.174 for all 100 hospitals. The efficiency rankings have a reasonable level of dependability since the standard errors are consistent across hospitals (SD of Bootstrap_SD = 0.002), indicating that estimation uncertainty is evenly spread throughout the panel rather than concentrated among particular hospitals. Fig. 6 shows bootstrap-based 95% confidence intervals for Dynamic SBM efficiency scores across hospitals.

FIGURE 6 BOOTSTRAP CONFIDENCE INTERVALS FOR DYNAMIC_SBM_EFFICIENCY



The composite score (Table 9) combines three different aspects of hospital performance into a single ranking measure. The formula assigns 50% weight to DEA efficiency, 25% to Markov persistence, and 25% to bootstrap stability. The final enhanced ranking, which is based on a composite score that combines Markov Persistence and Dynamic SBM Efficiency, shows the top 20 hospitals in Table 8 below.

TABLE 9 FINAL ENHANCED RANKINGS

Hospital_ID	Dynamic_SBM_Efficiency	Ei_norm	Mi	Bi	Ci	Final_Rank
H039	0.756107	1	0.818182	1	0.954545	1
H016	0.706908	0.789406	0.818182	0.147805	0.6362	2
H009	0.712335	0.812636	0.818182	0	0.610863	3
H022	0.64497	0.524283	0.818182	0.538843	0.601398	4
H084	0.555012	0.139222	1	0.810166	0.522152	5
H013	0.541276	0.080425	1	0.811694	0.493136	6
H037	0.563658	0.176231	0.693182	0.766574	0.453054	7
H070	0.522487	0	0.897727	0.890051	0.446945	8
H077	0.534368	0.050856	0.772727	0.825385	0.424956	9
H062	0.622437	0.427832	0.511364	0.229523	0.399137	10
H091	0.583404	0.260753	0.693182	0.281566	0.374063	11
H063	0.626936	0.447089	0.215909	0.169953	0.32001	12
H072	0.611881	0.382647	0	0.495514	0.315202	13
H050	0.580761	0.249439	0.261364	0.358416	0.279665	14
H081	0.626666	0.445934	0	0.17559	0.266864	15
H048	0.581759	0.253711	0.159091	0.335029	0.250385	16
H057	0.594665	0.308955	0.125	0.255088	0.249499	17
H040	0.590096	0.289397	0.136364	0.207091	0.230562	18
H021	0.584846	0.266925	0.136364	0.251162	0.230344	19
H052	0.589282	0.285913	0.125	0.206889	0.225929	20

Kendall's tau and Spearman's rank correlation coefficient were used to compare the final ranks with those from four different weighting situations. From Table 10, Kendall's tau values varied from 0.716 to 0.926, and Spearman correlation coefficients ranged from 0.881 to 0.983. These high correlation values show that there is a high degree of agreement between the baseline and alternative rankings, indicating that modest adjustments to the weighting structure do not significantly affect the ranking results. Even though there were minor differences in each hospital's position, the general ranking pattern was essentially the same in all cases. As a result, the suggested composite ranking approach can be regarded as reliable and consistent, offering assurance that the identification of high-performing institutions is independent of a particular weight selection.

TABLE 10 SENSITIVITY ANALYSIS

Scenario	Spearman_Rho	Kendall_Tau
Base	1	1
S1	0.977444	0.894737
S2	0.983459	0.926316
S3	0.914286	0.768421
S4	0.881203	0.715789

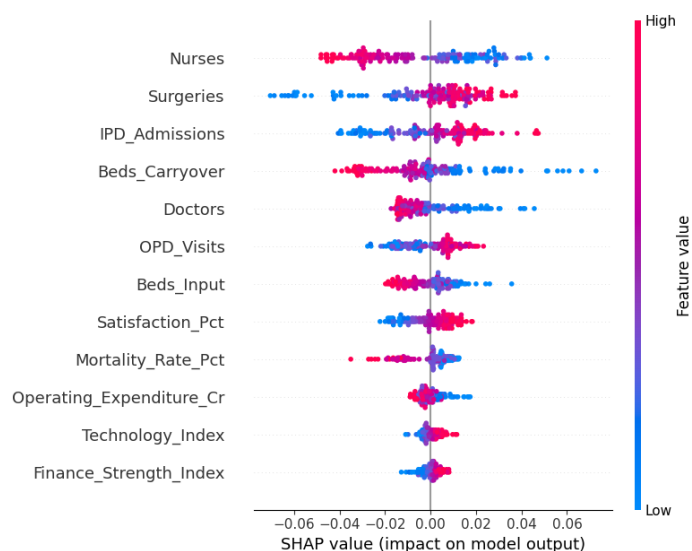
Sensitivity analysis and rank correlation metrics (Table 9) were used to further evaluate the robustness of ranking discrepancies. Strong agreement between the baseline and alternative ranking situations was shown by Spearman correlation coefficients ranging from 0.881 to 0.983 and Kendall's tau values ranging from 0.716 to 0.926. Further proof that

the ranking results are statistically sound and not influenced by small changes in model assumptions comes from these data, which show that the relative ordering of hospitals stayed mostly constant under various weighting systems.

ML MODEL VALIDATION RESULTS

Using 100 decision trees, the Random Forest model was built with a minimum sample split of two, a minimum leaf size of one, and no restrictions on the maximum tree depth. By using 100 trees, the model was able to reduce prediction variance through ensemble averaging and capture intricate nonlinear relationships between hospital performance variables. The trees were able to develop to their full potential because there was no maximum depth constraint, which improved the model's capacity to fit complex patterns in the training set. The discrepancy between the training R^2 (0.9407) and testing R^2 (0.4765), however, suggests that this flexibility may potentially lead to overfitting. Despite this, the cross-validated performance ($R^2 = 0.553 \pm 0.048$) shows that the model retains reasonable stability and considerable predictive capability across various data divisions. Overall, the chosen Random Forest design strikes a balance between model robustness and predictive accuracy, supporting its use in determining the major factors influencing hospital efficiency and producing projections for future efficiency. The stability and factors influencing hospital efficiency rankings derived from the Dynamic SBM-DEA analysis in conjunction with machine learning and bootstrap methods are summarized in Fig. 7.

FIG. 7 SHAP-BASED IDENTIFICATION OF KEY DRIVERS INFLUENCING HOSPITAL EFFICIENCY RANKINGS



The model obtained an R^2 of 0.4765, an RMSE of 0.5177, and an MAE of 0.4121 on the independent test dataset. According to these findings, the model maintained moderate prediction error levels while explaining roughly 47.65% of the variation in hospital efficiency. A comparison of the testing R^2 (0.4765) and training R^2 (0.9407) indicates that there may be some overfitting, but the consistent cross-validation performance and positive test R^2 show that the model still has significant predictive power on unknown data. Overall, the findings show (Table 11) a moderate level of out-of-sample predictive accuracy and offer enough proof to support hospital efficiency forecasting using the Random Forest model and subsequent SHAP-based interpretation.

TABLE 11 MACHINE LEARNING VALIDATION RESULTS

Metric	Value
Training Samples	640
Testing Samples	160
Training R^2	0.9407
Testing R^2	0.4765
CV R^2 (Mean $\pm$ SD)	0.553 $\pm$ 0.048

Metric	Value
RMSE	0.5177
MAE	0.4121

BAYESIAN FORECASTING RESULTS

The Bayesian forecasting results provide probabilistic estimates of future hospital efficiency and ranking positions, enabling a forward-looking assessment of performance. Results are given in Table 12 below.

TABLE 12 FORECASTED RANKING OF TOP 20 HOSPITALS

Hospital_ID	Future_Efficiency	Future_Rank
H009	0.689841	1
H039	0.689448	2
H016	0.661782	3
H022	0.629618	4
H062	0.614497	5
H063	0.611187	6
H081	0.610177	7
H072	0.602506	8
H052	0.587584	9
H040	0.586507	10
H057	0.583698	11
H048	0.581608	12
H091	0.578636	13
H050	0.577674	14
H021	0.574417	15
H013	0.539037	16
H037	0.536967	17
H077	0.531657	18
H084	0.530406	19
H070	0.519761	20

The present hospital rankings and the projected future rankings derived from the Bayesian fusion framework are contrasted in the accompanying table. The Rank_Change = Current Rank – Future Rank column shows whether a hospital's status is predicted to improve or deteriorate in the future. Table 13 rank transitions for the top 20 hospitals were predicted using Bayesian efficiency fusion (one period ahead). The Dynamic SBM Composite Score (CSi) provides the basis for Current_Rank. The posterior mean projection is known as the forecasted rank. A projected rank decline is indicated by a negative Rank_Change, whereas an improvement is shown by a positive Rank_Change.

TABLE 13 RANK TRANSITION

Hospital Rank	Hospital_ID	Current_Rank	Future_Rank(Forecasted)	Rank_Change
5	H084	5	19	-14
8	H070	8	20	-12
7	H037	7	17	-10
6	H013	6	16	-10
9	H077	9	18	-9
11	H091	11	13	-2
2	H016	2	3	-1
1	H039	1	2	-1

14	H050	14	14	0
4	H022	4	4	0
3	H009	3	1	2
16	H048	16	12	4
19	H021	19	15	4
13	H072	13	8	5
10	H062	10	5	5
12	H063	12	6	6
17	H057	17	11	6
15	H081	15	7	8
18	H040	18	10	8
20	H052	20	9	11

Each hospital's posterior credible intervals, rank confidence intervals, and Top 5 ranking probabilities were assessed to quantify the uncertainty in the Bayesian future efficiency forecasts and displayed in Table 14. The uncertainty surrounding the Bayesian posterior efficiency estimates after integrating data from the Random Forest forecasts and the DEA bootstrap results is represented by the value Posterior SD = 0.0039. So, the Bayesian fusion framework demonstrated minimal uncertainty in the projected future efficiency. As a result, the 95% Bayesian credible intervals surrounding the posterior efficiency estimates were small, indicating that the predictions are reliable and statistically accurate. Monte Carlo simulation was carried out by periodically sampling from the posterior efficiency distributions and recalculating hospital ranks to further evaluate ranking uncertainty.

TABLE 14 PROBABILITY OF SUSTAINING FUTURE RANKING

Hospital_ID	Posterior_Efficiency	Rank_Mean	Rank_CI_Lower	Rank_CI_Upper	Top5_Probability	Rank_Change
H009	0.689841	1.45405	1	2	1	2
H039	0.689448	1.54595	1	2	1	-1
H016	0.661782	3	3	3	1	-1
H022	0.629618	4.00004	4	4	1	0
H062	0.614497	5.26899	5	7	0.77853	5

The range of ranks that any hospital may reasonably occupy was shown by the rank confidence intervals that were produced, which offered a probabilistic measure of ranking variability. Furthermore, using the simulated rank distributions, the likelihood of being placed in the top five hospitals was calculated. Based on the comparison between current and forecasted rankings, H009, H039, H016, H022, and H062 can be identified as the most sustainable hospitals.

DISCUSSION

PRINCIPAL FINDINGS

This study combined Dynamic SBM DEA, bootstrap validation, Markov chain analysis, machine learning prediction, SHAP explainability, and Bayesian forecasting to create an integrated framework for hospital performance assessment and forecasting. Both consistently high-performing hospitals and those with substantial development potential were identified by the Dynamic SBM DEA results, which showed a great deal of variation in hospital efficiency. The stability of efficiency estimates was validated by bootstrap and Markov chain analyses, which also shed light on the persistence and evolution of hospital performance states across time. The framework produced probabilistic future rankings and identified hospitals likely to maintain or alter their performance positions by combining efficiency, stability, and persistence metrics with Random Forest and Bayesian forecasting.

COMPARATIVE STUDY WITH PREVIOUS WORK

In order to assess efficiency at a particular moment in time, prior hospital benchmarking studies have mostly relied on traditional DEA models. Although these methods offer useful information on relative performance, they frequently overlook ranking stability, temporal dynamics, uncertainty, and future performance expectations. Although dynamic DEA models are being used more and more to capture intertemporal interactions, the majority of research only concentrates on efficiency estimation and do not include uncertainty quantification or predictive analytics. By combining Dynamic SBM DEA with bootstrap validation and Markov chain analysis, the current study expands on previous research by addressing both statistical reliability and longitudinal performance transitions. Additionally, operational, clinical, and infrastructure-related efficiency factors can be identified by combining Random Forest regression and SHAP analysis, and Bayesian forecasting offers probabilistic future performance estimates. Therefore, compared to conventional DEA-based benchmarking frameworks, this integrated methodology represents a substantial methodological leap.

THEORETICAL IMPLICATIONS

By highlighting the benefits of integrating operational research techniques with machine learning and Bayesian inference, the work adds to the body of knowledge on hospital efficiency. The suggested framework takes a dynamic and predictive approach to hospital performance evaluation, going beyond static efficiency measurement. The paradigm offers a more thorough understanding of efficiency evolution over time by combining Bayesian updating, Markov persistence analysis, and bootstrap-based uncertainty evaluation. Although the proposed model demonstrated moderate predictive performance and satisfactory cross-validation stability, its robustness may be further improved through the use of real-world longitudinal hospital datasets. The present study relies on synthetically generated data, which may not fully capture the complexity and variability of actual healthcare environments. Future research incorporating real-time operational, financial, and clinical data may enhance model generalization, improve forecasting accuracy, and strengthen the reliability of machine-learning-based hospital efficiency prediction. The results bolster the claim that hospital performance should not be seen as a static efficiency score but rather as a dynamic process impacted by past performance trends, operational traits, and future uncertainty. The theoretical connection between efficiency assessment and data-driven decision support is further strengthened by the incorporation of explainable artificial intelligence through SHAP analysis.

MANAGERIAL AND POLICY IMPLICATIONS

By facilitating hospital benchmarking, identifying performance gaps, and evaluating performance persistence across time, the suggested approach offers healthcare managers and policymakers useful decision-support tools. By identifying the critical elements affecting efficiency, SHAP-based analysis enables focused enhancements in infrastructure, staffing, patient happiness, and service use. Furthermore, Bayesian forecasting provides probabilistic future rankings with uncertainty measurements, making it easier to make well-informed judgments on capacity planning, long-term performance management, and resource allocation. Furthermore, by identifying hospitals that are anticipated to perform better or worse in the future, the predicted ranking changes offer important managerial and policy insights. While hospitals with notable rank increases serve as developing benchmarks for performance improvement and best-practice adoption, hospitals predicted to have major rank reductions may need proactive operational assessment and resource optimization to minimize future efficiency losses. Additionally, by supplementing conventional cross-sectional assessments with a forward-looking evaluation of performance sustainability, these forecasts can assist accreditation agencies and healthcare regulators, turning efficiency measurement into a dynamic tool for quality monitoring and strategic decision-making.

LIMITATIONS AND FUTURE SCOPE

The following constraints should be noted even if the simulated dataset shows good face validity and statistical agreement with Indian multispecialty hospital benchmarks:

- The interpretation of surgical volumes and fatality rates in highly specialized centers may be impacted by the dataset's failure to capture specialty-level disaggregation (e.g., cardiology vs. oncology numbers). The dataset is primarily typical of pure private multispecialty hospitals.
- Teaching hospital characteristics (resident physicians, research output) are not explicitly modeled. Although the range of hospital sizes implicitly incorporates this diversity, geographic heterogeneity (Tier-1 vs. Tier-2 cities, North vs.

South India) is not explicitly classified. The difference between IT infrastructure, buildings, and medical equipment is not made; instead, capital stock is modeled as an aggregate carry-over variable.

- Although the system allows evidence-based benchmarking and policy decisions, robustness may be impacted by simulated data, assumptions in priors, and RF structure [27,28].
- The actual multivariate correlation structure of the original data is not explicitly preserved in the synthetic dataset. Rather, temporal carryover rules and structural factors at the hospital level create inter-variable dependence.

In order to properly capture measurement error, operational variability, and institutional variation, future research should evaluate the suggested approach using real-world longitudinal hospital datasets. To specifically maintain the original data's empirical multivariate correlation structure, Gaussian Copula-based simulation or other dependency-preserving techniques to produce synthetic data that more closely mimics the observed multivariate dependence; future research may be able to get around this restriction.

To enhance the depiction of clinical complexity, additional quality-of-care indicators, case-mix modifications, and specialty-specific performance metrics may be added. Furthermore, to improve uncertainty modelling and forecast accuracy in hospital performance evaluation, sophisticated techniques such as stochastic DEA, fuzzy DEA, neutrosophic DEA, and Bayesian hierarchical models could be investigated. Numerous service sectors, such as dining establishments, lodging facilities, educational institutions, banking, and transportation, can use the suggested approach. It is especially helpful in settings where resource allocation and strategic planning depend on precise efficiency evaluation, strong ranking systems, uncertainty management, and trustworthy performance predictions.

CONCLUSION

The framework provides more trustworthy insights into hospital performance by producing probabilistic efficiency rankings while considering intertemporal slack uncertainty. The findings show that Bayesian fusion increases predictive accuracy with quantifiable ranking confidence and that slack-based factors have a major impact on efficiency. By assisting administrators in differentiating between fundamental resource limitations and operational inefficiencies, the method will promote well-informed administrative choices. Overall, the integrated framework is useful in any service industry like restaurants, hotels, educational institutes, banking sectors, transportation, etc where strategic planning and resource optimization depend on efficiency measurement, ranking stability, controlling uncertainty, and forecasting future performance.

References

1. Persis DJ, S A, Sunder M V, G R, Sreedharan VR, Saikouk T. Improving patient care at a multi-speciality hospital using lean six sigma. *Production Planning & Control*. 2022 Sep 10;33(12):1135-54.
2. Chmielewska M, Stokwiszewski J, Markowska J, Hermanowski T. Evaluating organizational performance of public hospitals using the McKinsey 7-S framework. *BMC health services research*. 2022 Jan 2;22(1):7.
3. Saha E, Rathore P. Case mix-based resource allocation under uncertainty in hospitals: Physicians being the scarce resource. *Computers & Industrial Engineering*. 2022 Dec 1;174:108767.
4. Li Z, Feng C, Tang Y. Bank efficiency and failure prediction: a nonparametric and dynamic model based on data envelopment analysis. *Annals of Operations Research*. 2022 Aug;315(1):279-315.
5. Hajiagha SH, Amoozad Mahdiraji H, Hashemi SS, Garza-Reyes JA, Joshi R. Public hospitals performance measurement through a three-staged data envelopment analysis approach: evidence from an emerging economy. *Cybernetics and Systems*. 2023 Jan 2;54(1):1-26.
6. Nguyen PA, Tran BL, Simioni M. Optimal capital adequacy ratio: An investigation of Vietnamese commercial banks using two-stage DEA. *Cogent business & management*. 2021 Jan 1;8(1):1870796.
7. Zervopoulos PD, Kanas A, Fouteris A, Cheng G, Alqasimi SA. A Bayesian approach for correcting bias of data envelopment analysis estimators using the super-efficiency frontier. *Socio-Economic Planning Sciences*. 2025 Aug 6:102299.

8. Nagy G. Performance Measurement and Quality Assurance in Higher Education: Application of DEA, AHP, and Bayesian Models. *Trends in Higher Education*. 2025 Sep 18;4(3):54.
9. Kawaguchi H, Tone K, Tsutsui M. Estimation of the efficiency of Japanese hospitals using a dynamic and network data envelopment analysis model. *Health care management science*. 2014 Jun;17(2):101-12.
10. Khushalani J, Ozcan YA. Are hospitals producing quality care efficiently? An analysis using Dynamic Network Data Envelopment Analysis (DEA). *Socio-economic planning sciences*. 2017 Dec 1;60:15-23.
11. Gorgemans S, Comendeiro-Maaløe M, Ridao-López M, Bernal-Delgado E. Quality and technical efficiency do not evolve hand in hand in Spanish hospitals: Observational study with administrative data. *PLoS One*. 2018 Aug 2;13(8):e0201466.
12. Kočiřová K, Sopko J. The Efficiency of Public Health and Medical Care Systems in EU Countries: Dynamic Network Data Envelopment Analysis. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. 2020 Mar 1;68(2).
13. Zhang T, Lu W, Tao H. Efficiency of health resource utilisation in primary-level maternal and child health hospitals in Shanxi Province, China: a bootstrapping data envelopment analysis and truncated regression approach. *BMC health services research*. 2020 Mar 6;20(1):179.
14. Gavurova B, Kocisova K, Sopko J. Health system efficiency in OECD countries: dynamic network DEA approach. *Health Economics Review*. 2021 Oct 12;11(1):40.
15. Mischczynska K, Mischczyński PM. Measuring the efficiency of the healthcare sector in Poland—a window-DEA evaluation. *International Journal of Productivity and Performance Management*. 2022 Aug 16;71(7):2743-70.
16. Peykani P, Memar-Masjed E, Arabjazi N, Mirmozaffari M. Dynamic performance assessment of hospitals by applying credibility-based fuzzy window data envelopment analysis. In *Healthcare* 2022 May 9 (Vol. 10, No. 5, p. 876). MDPI.
17. Mbau R, Musiega A, Nyawira L, Tsofa B, Mulwa A, Molyneux S, Maina I, Jemutai J, Normand C, Hanson K, Barasa E. Analysing the efficiency of health systems: a systematic review of the literature. *Applied health economics and health policy*. 2022 Dec 28;21(2):205.
18. El Husseiny IA. The efficiency of healthcare systems in the Arab countries: a two-stage data envelopment analysis approach. *Journal of Humanities and Applied Social Sciences*. 2023 Aug 28;5(4):339-58.
19. González-de-Julián S, Vivas-Consuelo D, Barrachina-Martínez I. Modelling efficiency in primary healthcare using the DEA methodology: an empirical analysis in a healthcare district. *BMC health services research*. 2024 Aug 25;24(1):982.
20. Afonso GP, Ferreira DC, Figueira JR. A Network-DEA model to evaluate the impact of quality and access on hospital performance. *Annals of Operations Research*. 2024 Nov;342(3):2169-99.
21. Koçak E, Gökgöz F. Dynamic Network DEA Approach for Evaluating Hospital Performance in the Healthcare Sector. *International Journal of Advances in Engineering and Pure Sciences*. 2025 Sep 9;37(3):252-62.
22. Van Breugel B, Qian Z, Van Der Schaar M. Synthetic data, real errors: how (not) to publish and use synthetic data. In *International Conference on Machine Learning* 2023 Jul 3 (pp. 34793-34808). PMLR.
23. Tone K, Tsutsui M. Dynamic DEA: A slacks-based measure approach. *Omega*. 2010 Jun 1;38(3-4):145-56.
24. Saen RF, Yousefi F, Azadi M. Artificial intelligence powered predictions: enhancing supply chain sustainability. *Annals of Operations Research*. 2024 Jun 15:1-44.
25. Al Harbi S, Aljohani B, Elmasry L, Baldovino FL, Raviz KB, Altowairqi L, Alshlowi S. Streamlining patient flow and enhancing operational efficiency through case management implementation. *BMJ open quality*. 2024 Feb 28;13(1).
26. Zhao Y. Sustainable Performance Evaluation and Prediction of the Banking Sector: Opening the Black Box of DEA with Machine Learning and Explainable AI. In *International Conference on Machine Learning and Computing* 2025 Feb 14 (pp. 795-809). Cham: Springer Nature Switzerland.
27. Kohl, S., Schoenfelder, J., Fügner, A., & Brunner, J. O. (2019). The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals. *Health care management science*, 22(2), 245-286.
28. Ferreira, D. C., Nunes, A. M., & Marques, R. C. (2020). Operational efficiency vs clinical safety, care appropriateness, timeliness, and access to health care: The case of Portuguese public hospitals. *Journal of Productivity Analysis*, 53(3), 355-375.
29. National Accreditation Board for Hospitals and Healthcare Providers (NABH). *NABH Standards for Hospitals*, 5th Edition. Quality Council of India, 2020.