Research Annals of Industrial and Systems Engineering

www.raise.reapress.com

Res. Ann. Ind. Syst. Eng. Vol. 2, No. 4 (2025) 238–248.

Paper Type: Original Article

Value Efficiency in Centralized Resource Allocation with Missing Data: A Case Study of Construction Projects

Ali Sedaghatpour[1,*], Morteza Shafiee[2], Mohammad Reza Mozaffari[3], Hassan Soltani[2]

[1] Department of Industrial Management, Kish International Branch, Islamic Azad University, Kish, Iran; a.sedaghatpour@iau.ac.ir.
[2] Department of Industrial Management, Shiraz Branch, Islamic Azad University, Shiraz, Iran; morteza.shafiee@iau.ac.ir; h.soltani@iau.ac.ir.
[3] Department of Mathematics, Shiraz Branch, Islamic Azad University, Shiraz, Iran; mr.mozaffari@iau.ac.ir.

Citation:

Received: 25 May 2025 *Revised: 01 July 2025* *Accepted: 12 October 2025*	Sedaghatpour, A., Shafiee, M., Mozaffari, M. R., & Soltani, H. (2025). Value efficiency in centralized resource allocation with missing data: A case study of construction projects. *Research annals of industrial and systems engineering*, 2(4), 238-248.

Abstract

This study proposes a mathematical model for evaluating the performance of construction projects based on the Value Efficiency (EV) approach under conditions of missing data. The proposed framework integrates Data Envelopment Analysis (DEA) with missing data estimation techniques to assess the relative efficiency of Decision-Making Units (DMUs) in real-world environments characterized by incomplete information. To address missing values in undesirable outputs, two replacement strategies are employed: mean substitution and an efficiency-based scenario approach. The empirical application includes 33 construction projects affiliated with Sadra New Town Development Company in Shiraz (before 2024). The results demonstrate that the proposed model effectively distinguishes between efficient and inefficient projects and provides a scientific basis for improving Centralized Resource Allocation (CRA) decisions.

Keywords: Value efficiency, Data envelopment analysis, Missing data, Centralized resource allocation, Construction projects.

1 | Introduction

Performance evaluation and resource allocation in multi-project environments have long been recognised as critical drivers of organisational efficiency, particularly within capital-intensive industries such as construction. In such settings, a central decision-maker typically oversees a portfolio of interdependent projects, each consuming shared inputs such as financial capital, labour, and time, while generating outputs that differ markedly in nature, magnitude, and measurability. The challenge is compounded by the simultaneous presence

Corresponding Author: a.sedaghatpour@iau.ac.ir
https://doi.org/10.22105/raise.v2i4.92

of desirable outputs, which managers seek to maximise (e.g., physical progress, quality scores, and net payments), and undesirable outputs, which they seek to minimise (e.g., construction waste, rework costs, and environmental residuals). Designing a rigorous, systematic framework for evaluating and improving the performance of such a heterogeneous project portfolio under realistic data conditions represents both a theoretical and practical imperative.

Data Envelopment Analysis (DEA), pioneered by Charnes et al. [1] in their seminal work on measuring the efficiency of Decision-Making Units (DMUs), has emerged as the dominant non-parametric methodology for multi-input, multi-output performance benchmarking. Since its introduction, DEA has undergone extensive theoretical elaboration, most notably the incorporation of Variable Returns to Scale (VRS) by Banker et al. [2], and has been deployed across a broad spectrum of domains, including banking, healthcare, education, transportation, and construction project management [3], [4]. Within the construction sector, DEA has demonstrated considerable utility in revealing performance differentials among projects or firms and in providing actionable input-output targets that inform managerial decision-making [5], [6].

A significant limitation of classical DEA models, however, is their fundamentally decentralised orientation. Each DMU is evaluated independently, and efficiency targets are set without regard to the system-level consequences of resource reallocation. In organisational contexts characterised by a single governing authority, such as urban development corporations, public works agencies, or project management offices supervising a portfolio of construction contracts, this unit-by-unit perspective is suboptimal. To address this limitation, Lozano and Villa [7] introduced the Centralised Resource Allocation (CRA) framework, which reframes the objective from the measurement of individual DMU efficiency to the simultaneous optimisation of aggregate input consumption or output production across all units. Unlike conventional DEA, CRA models determine practical input and output targets for every DMU under a shared technology, making them particularly suited for centrally managed portfolios. Subsequent extensions have refined this approach under a variety of assumptions, including VRS conditions (Asmild et al. [8], cost-revenue considerations Mar-Molinero et al. [9], and network structures Fang [10]).

A further dimension of complexity arises from the treatment of undesirable outputs within DEA-based efficiency models. In many real-world production processes, and construction projects in particular, certain outputs are inherently negative externalities: construction waste, safety incidents, or cost overruns represent outcomes that are not voluntarily produced and whose magnitude is partially governed by environmental or contractual factors beyond managerial discretion. The formal incorporation of undesirable outputs into DEA models has been addressed through several competing approaches, including monotone-decreasing data transformations (Seiford and Zhu [11], weak disposability constraints [12], and directional distance functions [13]). Each approach carries distinct technical implications for the production possibility set and for the resulting efficiency scores. Within the CRA literature, the integration of undesirable outputs adds a further layer of methodological complexity that has received comparatively limited attention.

Perhaps the most practically consequential yet methodologically underexplored challenge in operational DEA applications is the presence of missing data. Real-world datasets, particularly those derived from longitudinal construction project records, routinely exhibit incomplete observations. Missing values may arise from a variety of operational causes: delayed reporting of waste or residual outputs in long-duration contracts, strategic withholding of sensitive performance indicators for post-project litigation purposes, or simple administrative omission. Standard DEA formulations require complete data matrices; the conventional response of deleting DMUs with missing values distorts the composition of the reference set and biases the efficiency scores of the remaining units [14], [15]. More principled imputation strategies, including mean substitution, fuzzy set approaches, interval-valued DEA, and hot-deck methods, have been proposed in the literature [15], [16]. Yet, their integration with CRA models and undesirable output frameworks remains largely unexplored.

This paper addresses the intersection of these three methodological threads, CRA, undesirable outputs, and missing data, by proposing a novel mathematical framework that integrates the Efficiency Value (EV) approach with DEA under conditions of incomplete information. EV, introduced by Halme et al. [17] within a DEA context, allows a preferred solution or Most Productive Scale (MPS) to guide the efficiency assessment, thereby incorporating decision-maker preferences directly into the target-setting process. Building on this foundation, the present study develops a CRA model that accommodates undesirable outputs governed by equality constraints, reflecting their exogenous, partially uncontrollable nature, and introduces two complementary strategies for imputing missing values in those outputs. The first strategy employs mean substitution across available observations, offering simplicity and transparency. The second adopts an efficiency-based scenario approach, in which the missing output value is determined by each DMU's efficiency status computed on the reduced dataset: efficient units retain values inferred from their reference set, while inefficient units receive the mean of observed peers. This dual-strategy design allows practitioners to assess the sensitivity of allocation targets to the imputation methodology.

The empirical grounding of the proposed framework is provided through a case study of 33 construction projects managed by the Sadra New Town Development Company in Shiraz, Iran. The dataset is characterised by substantial heterogeneity in contract duration and financial scale, by the presence of both desirable outputs (net payment, quality percentage, physical progress) and an undesirable output (construction waste), and by partially missing waste observations, a configuration that renders it an ideal testbed for the proposed methodology. To the best of the authors' knowledge, no prior study has simultaneously extended the CRA framework to accommodate undesirable outputs under VRS technology while embedding principled missing-data imputation within the target-setting procedure.

The contributions of this paper are threefold: 1) it advances the CRA literature by incorporating undesirable outputs into the centralised target-setting framework under VRS, offering a more realistic representation of construction project technology, 2) it bridges the gap between missing data imputation and CRA by developing two structured replacement strategies that preserve the theoretical coherence of the production possibility set, and 3) it demonstrates the practical applicability of the proposed models through a real-world empirical application, offering decision-relevant insights for construction management in new urban development contexts.

2 | Fundamentals of Data Envelopment Analysis

In the framework of DEA, when outputs are classified into desirable and undesirable categories, the production possibility set is defined based on an m-dimensional input vector and two distinct output vectors. The underlying technology is constructed through convex combinations of the observed DMUs. In this setting, for a given input vector $X_j = (x_{1j}, x_{2j}, \dots, x_{mj})$, the desirable outputs $Y_j = (y_{1j}, y_{2j}, \dots, y_{sj})$ are subject to inequality constraints and may expand up to the level determined by the reference convex combination, reflecting their controllable nature under managerial discretion. In contrast, the undesirable outputs $\acute{Y}_j = (y_{1j}, y_{2j}, \dots, y_{sj})$ are incorporated through equality constraints, requiring them to remain fixed at the level of the corresponding convex combination, since these outputs are determined by exogenous or environmental factors beyond managerial control. Consequently, the modified production possibility set forms a conditional convex hull that preserves convexity and non-negativity assumptions while enabling a more equitable efficiency assessment by disentangling managerial performance from environmental influences.

Centralized Resource Allocation (CRA) models were proposed by Lozano and Villa [7]. These models evaluate DMUs with the aim of minimizing total inputs and maximizing total outputs. Unlike traditional DEA models, the primary objective of CRA models is not to measure efficiency, but rather to determine target levels for each DMU.

$$
\begin{aligned}
&\sigma^* = \text{Min}\ \sigma \\
&\text{s.t.} \quad \sum_{t=1}^{n}\sum_{j=1}^{n} \beta_{tj}\, x_{ij} \le \sigma\left(\sum_{j=1}^{n} x_{ij}\right), \forall i, \\
&\qquad \sum_{t=1}^{n}\sum_{j=1}^{n} \beta_{tj}\, y_{rj} \ge \left(\sum_{j=1}^{n} y_{rj}\right), r \in O_1, \\
&\qquad \sum_{j=1}^{n} \beta_{tj} = 1, \forall t, \\
&\qquad \beta_{tj} \ge 0.
\end{aligned}
\tag{1}
$$

Model (1) represents an input-oriented CRA framework. The objective is to minimize the proportional scaling factor σ applied to the inputs of a DMU, while ensuring that the desirable outputs at least reach their current observed levels. The model constructs a convex combination of observed units using intensity variables β_{tj}, subject to convexity constraints ($\sum_{j=1}^{n} \beta_{tj} = 1$) and non-negativity ($\beta_{tj} \ge 0$). For each input i, the weighted sum of reference inputs is constrained to be no greater than σ times the observed input, reflecting the goal of input reduction. Simultaneously, for each desirable output $r \in O_1$, the weighted sum must be at least equal to the observed output, representing output preservation or improvement. This formulation allows the model to determine target input levels for each DMU under a technology with VRS, without directly measuring efficiency, in line with the objectives of CRA models.

Model (2) represents the target-setting stage based on the optimal solution obtained from *Model (1)*. It determines the target input and desirable output levels for each DMU by forming a convex combination of reference units. This approach allows each unit to identify practical resource allocation targets while respecting the technology's convexity and VRS assumptions.

$$
\left(\sum_{j=1}^{n} \beta_{tj}^{*}\, x_{ij},\ \sum_{j=1}^{n} \beta_{tj}^{*}\, y_{rj}\right), \forall t. \tag{2}
$$

3 | Modeling

In this section, by introducing the production possibility set in CRA models and by considering missing data in the undesirable outputs of DMUs, the production possibility set under VRS technology is defined as follows:

$$P_1 = \left\{ (A, B, D) \mid \begin{array}{c} \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} X_j \leq 1A, \\ \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} Y_j \geq 1B, \\ \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} \acute{Y}_j = 1D, \\ \sum_{j=1}^{n} \beta_{tj} = 1, \beta_{tj} \geq 0, \forall t, \forall j \end{array} \right\}.$$

The production possibility set is defined by considering the inclusion of all observations, encompassing the inputs as well as the desirable and undesirable outputs. According to the possibility principle, higher levels of inputs are permissible within the production set, whereas lower levels of desirable outputs correspond to feasible production. In contrast, undesirable outputs, which may include missing data, do not necessarily satisfy the possibility principle. Furthermore, in order to comply with the assumptions of VRS and convexity, convexity constraints are imposed on the intensity variables corresponding to both inputs and outputs. Collectively, these considerations ensure that the production possibility set accurately reflects the technological capabilities while differentiating between desirable and undesirable outputs and accounting for data limitations.

$$\begin{aligned} \alpha^* = \text{Min} \quad & \left\{ \textstyle\sum_{i=1}^{m} \rho_i - \sum_{r \in O_1} \varphi_r \right\}, \\ \text{s.t.} \quad & \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} x_{ij} \leq \rho_i \left(\sum_{j=1}^{n} x_{ij} \right), \forall i, \\ & \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} y_{rj} \geq \varphi_r \left(\sum_{j=1}^{n} y_{rj} \right), r \in O_1, \\ & \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} \acute{y}_{rj} = \left(\sum_{j=1}^{n} \acute{y}_{rj} \right), r \in O_2, \\ & \sum_{j=1}^{n} \beta_{tj} = 1, \forall t, \\ & \beta_{tj}: \begin{cases} \geq 0 & \text{if } \beta_{tj}^* = 0, \\ \text{free} & \text{if } \beta_{tj}^* > 0. \end{cases} \end{aligned} \tag{3}$$

Since $\sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj} \acute{y}_{rj} = \left(\sum_{j=1}^{n} \acute{y}_{rj} \right)$ corresponds to missing data in the undesirable outputs, two potential approaches can be applied to address this issue, which are discussed below. In general, missing data arise in situations where either the data do not exist, or the decision-maker is unable to include them in the evaluation. For example, in long-term projects, waste or residuals may be both unavoidable and financially significant, and managers may retain this information for addressing potential problems after project completion.

Therefore, the simplest approach for handling missing data in undesirable outputs is to replace them with the mean value of that output across the available observations, which is straightforward and intuitive. In this approach, the mean is calculated for each undesirable output separately and then used as a replacement.

The second approach is slightly more sophisticated. First, the undesirable output is removed from all DMUs, and the efficiency of each unit is calculated based on the reduced set of outputs. If a unit remains efficient after the removal, the missing undesirable output is replaced by the values derived from this efficient scenario. If a unit becomes inefficient, the missing undesirable output is replaced by the mean value of that output across all units. This strategy effectively trusts the reference units for efficient DMUs and incorporates the mean of the reference undesirable outputs into the evaluation. For inefficient units, using the overall mean of all units' undesirable outputs provides a more reasonable and consistent replacement. In general, missing data are substituted with the average values of the relevant undesirable outputs from the reference or observed units.

Therefore, the proposed model can be implemented as follows, considering missing data under the second strategy, which involves the evaluation of the unit under assessment. Since CRA models do not compute efficiency but rather determine targets, this approach provides the opportunity to verify whether the computed targets are appropriate for the evaluated unit. The advantage of this perspective is that there is no need to solve a separate linear programming model, as the assessment and comparison can be performed directly based on the derived targets.

$$
\begin{aligned}
Z_{\hat{r}}^{*} = \text{Min} \quad & \left\{\sum_{i=1}^{m} \rho_i - \sum_{r\in O_1} \varphi_r\right\}, \\
\text{s.t.} \quad & \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj}\, x_{ij} \le \rho_i \left(\sum_{j=1}^{n} x_{ij}\right), \forall i, \\
& \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj}\, y_{rj} \ge \varphi_r \left(\sum_{j=1}^{n} y_{rj}\right), r \in O_1, \\
& \sum_{t=1}^{n} \sum_{j=1}^{n} \beta_{tj}\, \acute{y}_{\hat{r}j} = \left(\sum_{j=1}^{n} \acute{y}_{\hat{r}j}\right), \hat{r} \in O_2, \\
& \sum_{j=1}^{n} \beta_{tj} = 1, \forall t, \\
& \beta_{tj}: \begin{cases} \ge 0 & \text{if } \beta_{tj}^{*} = 0, \\ \text{free} & \text{if } \beta_{tj}^{*} > 0. \end{cases}
\end{aligned} \tag{4}
$$

$Z_{\hat{r}}^{*}$ represents the optimal value of *Model (4)* obtained after the removal of the missing undesirable output.

4 | Case Study: Evaluation of Construction Projects

The dataset used in this study corresponds to construction projects in Shiraz during the period before 2024, encompassing 33 DMUs with varying contract durations and amounts. The input data, summarized in *Table 1*, include the duration of each contract in months and the total contract amount in Iranian Rials (IRR).

Contract durations range from 4 to 24 months, reflecting both short-term and long-term projects, while contract amounts vary widely from approximately 5 billion IRR to over 369 billion IRR, indicating substantial differences in project scale and financial magnitude. This heterogeneity provides a suitable foundation for efficiency analysis and resource allocation modeling, allowing for meaningful comparisons across contracts with diverse temporal and financial characteristics.

Table 1. Input data for construction projects in Shiraz.

DMU	Contract Duration (Months)	Contract Amount (IRR)
DMU1	15	214,658,003,006
DMU2	9	29,502,267,818
DMU3	4	20,890,000,000
DMU4	12	15,906,909,491
DMU5	12	208,428,788,230
DMU6	12	329,852,019,398
DMU7	24	369,567,223,439
DMU8	8	59,428,956,801
DMU9	12	178,376,124,198
DMU10	12	6,882,330,256
DMU11	9	103,116,611,117
DMU12	5	49,946,882,317
DMU13	5	5,473,576,812
DMU14	6	96,218,907,142
DMU15	9	53,871,359,908
DMU16	6	82,875,788,632
DMU17	8	59,428,956,801
DMU18	4	18,910,426,298
DMU19	5	49,946,882,317
DMU20	4	9,998,257,790
DMU21	12	295,974,083,027
DMU22	12	212,864,337,075
DMU23	12	22,499,969,841
DMU24	12	185,052,109,638
DMU25	12	21,543,869,808
DMU26	12	64,018,039,674
DMU27	6	19,862,692,357
DMU28	15	71,338,455,757
DMU29	12	152,101,736,374
DMU30	6	70,127,984,801
DMU31	6	139,468,014,734
DMU32	6	21,352,898,507
DMU33	7	22,162,094,183

The input data for the 33 construction projects in Shiraz (before 2024) exhibits substantial variability. Contract durations range from 4 to 24 months, with an average of approximately 10.4 months and a median of 12 months, reflecting a mix of short-term and long-term projects. Contract amounts span from approximately 5.47 billion IRR to 369.57 billion IRR, with a mean of around 104.5 billion IRR and a median of 59.43 billion IRR, indicating a wide distribution of project scales and financial magnitudes. This heterogeneity in both temporal and financial inputs provides a suitable foundation for efficiency analysis and resource allocation modeling, enabling meaningful comparisons across diverse DMUs.

Table 2 presents the output data for 33 construction projects in Shiraz, categorized into desirable and undesirable outputs. In this context, net payment (IRR), project quality percentage, and physical progress are considered desirable outputs, as higher values indicate better performance and managerial effectiveness. In contrast, construction waste (t) represents an undesirable output, where lower values are preferred, and some observations include missing data. This distinction allows efficiency analysis to reward good performance while accounting for outcomes that are detrimental or partially outside managerial control.

Table 2. Desirable and undesirable outputs for construction projects in Shiraz.

DMU	Net Payment (IRR)	Project Quality (%)	Physical Progress	Construction Waste (t)
DMU1	299,029,881,431	0.90	0.6101	23
DMU2	42,089,482,794	0.95	0.8042	12
DMU3	11,744,133,386	1.00	0.5000	34
DMU4	16,917,708,989	0.90	0.9500	18
DMU5	100,061,008,195	0.85	0.4681	Missing
DMU6	574,565,323,793	0.90	0.9904	40
DMU7	432,873,310,570	0.95	0.4500	15
DMU8	44,620,363,241	1.00	0.5000	22
DMU9	211,599,104,045	0.95	0.5200	19
DMU10	1,590,620,054	0.85	0.3470	28
DMU11	11,666,168,523	0.85	0.2700	35
DMU12	58,608,431,530	0.90	0.3250	21
DMU13	4,891,793,454	0.90	1.0000	17
DMU14	75,576,714,865	0.95	0.4454	25
DMU15	57,392,859,096	0.93	0.3692	30
DMU16	19,023,377,048	0.95	0.6937	Missing
DMU17	44,620,363,241	1.00	0.5000	24
DMU18	13,936,764,895	0.93	0.8636	13
DMU19	58,608,431,530	0.88	0.5100	37
DMU20	49,608,905,890	0.92	0.7099	29
DMU21	188,640,352,531	0.85	0.7110	26
DMU22	277,189,035,744	0.90	0.9049	41
DMU23	18,631,152,105	0.95	0.6413	14
DMU24	241,597,908,396	0.98	0.3305	33
DMU25	27,262,871,288	0.86	0.9541	20
DMU26	66,541,762,915	0.94	0.6479	38
DMU27	5,941,464,883	0.96	0.5000	Missing
DMU28	111,003,048,317	0.91	0.9984	23
DMU29	56,229,217,796	0.90	0.3000	31
DMU30	91,673,391,347	0.90	0.6088	22
DMU31	86,776,284,286	0.96	0.8012	28
DMU32	3,895,318,317	0.91	0.6730	Missing
DMU33	5,205,097,149	0.85	0.8235	27

The output variables show substantial variability across the 33 DMUs. Desirable outputs, including net payment (IRR), project quality percentage, and physical progress, vary widely: net payment ranges from 1.59 billion IRR to 574.57 billion IRR, project quality percentage ranges from 0.85 to 1, and physical progress ranges from 0.27 to 0.9984. The undesirable output, construction waste, ranges from 12 to 41 tons, with some missing data points requiring careful handling or imputation during analysis. The presence of missing values in undesirable outputs highlights the need for appropriate strategies to ensure fair efficiency assessment, while the variability in desirable outputs reflects differences in managerial performance and project execution outcomes.

By examining the θ_{CCR} and θ_{BCC} values, it is evident that several units exhibit full radial and non-radial efficiency (CCR = 1 and BCC = 1), such as units 3, 6, 10, 13, 18, 20, 27, and 32. These results suggest that these units lie on the efficiency frontier and demonstrate optimal performance relative to their reference sets. In contrast, units such as 2, 4, 5, 7, 9, and 29 exhibit lower efficiency levels (CCR < 0.6), implying that they are improvable and require reductions in inputs or increases in outputs. The comparison between CCR and BCC shows that for some units, such as 2 and 5, non-radial efficiency BCC is higher than radial efficiency, indicating that part of their inefficiency arises from scale effects.

The s^*_add values represent the additional number of inputs or the output shortfalls required to reach efficiency. Units with s_add values close to zero, such as units 2, 3, 5, 6, 10, 16, and 32, demonstrate that they can reach the efficiency frontier with minimal adjustments. Conversely, units such as 1, 9, 11, 21, and 29 exhibit substantially high s_add values, indicating that significant input reductions or output increases are necessary for these units to achieve optimal efficiency. Therefore, slack values serve as a critical indicator for identifying critical points and corrective needs in each unit.

The Eff_AP values, as indicators of overall aggregate efficiency, show that certain units, such as 6 (Eff_AP ≈ 2.07), 13 (Eff_AP ≈ 2.57), and 20 (Eff_AP ≈ 3.26), perform significantly better than others. These units efficiently convert inputs into outputs and surpass the average performance within the network. In contrast, units such as 23 (Eff_AP ≈ 0.38), 29 (Eff_AP ≈ 0.35), and 4 (Eff_AP ≈ 0.43) demonstrate very low performance, requiring managerial attention and intervention. This indicator is useful for identifying both benchmark and underperforming units.

The Eff_Maj and Eff_jlk values reflect partial efficiency regarding outputs and multi-dimensional overall efficiency, respectively. For most units, Eff_jlk equals 1.0, indicating that all units meet the minimum standard of efficiency when considering their reference sets. However, Eff_Maj values vary for units such as 7 (0.49), 29 (0.68), and 2 (0.94), indicating imbalances in output performance. This analysis suggests that certain units are inefficient in converting inputs into all outputs, highlighting the need for targeted improvement in specific output dimensions.

An examination of the reference set reveals that efficient units frequently serve as benchmarks for other units. For instance, units 6, 20, and 13 repeatedly appear as reference units, emphasizing their role in establishing performance standards across the network. This insight can guide efficiency improvement policies by facilitating knowledge transfer and best practices from high-performing units to underperforming ones. The recurring pattern of reference sets suggests that corrective focus should be directed toward low-performing units with multiple references (e.g., units 1, 2, 4, 29) to maximize resource utilization and overall network efficiency.

In *Model (4)*, when DMU6 is considered as the MPS, the efficiency analysis and target setting indicate that to achieve full efficiency, this unit needs to adjust its inputs and optimize its outputs. Specifically, the first input (contract duration) should be reduced to 9.8966 months, and the second input (contract amount) should be adjusted to 2.4737×10^{10} IRR, reflecting a significant reduction in resource consumption. Meanwhile, the outputs should be increased to net payment = 1.2274×10^{11} IRR, project quality = 2.2762 (on a normalized scale), and physical progress = 1.7564 to reach the efficiency frontier. These target values demonstrate that, through appropriate resource allocation and output enhancement, DMU6 can achieve both technical efficiency and scale efficiency, serving as a benchmark for other units.

5 | Conclusion

The proposed CRA model provides achievable input and output targets for each DMU, focusing on performance improvement and optimal resource utilization. This model can serve as an effective tool for decision-making in construction companies in new cities, allowing managers to define practical objectives for resource allocation and project performance enhancement. The proposed algorithm for imputing missing data enhances the accuracy and reliability of the analysis, increasing the practical applicability of the results in real-world decision-making.

For future research, it is recommended to employ Machine Learning (ML) techniques for estimating missing data to further improve target-setting accuracy. Additionally, extending the model into fuzzy or stochastic frameworks could allow better handling of uncertainty and environmental variability, enhancing the model's applicability in more complex conditions. This integration of resource optimization and uncertainty analysis offers a comprehensive framework for managing construction projects.

Authors' Contributions

The author solely conducted the research and prepared the manuscript and has approved its final version.

Funding

This work was carried out without financial support from any public, commercial, or non-profit organizations.

Data Availability

The data are available from the corresponding author upon reasonable request.

Conflict of Interest

There are no competing interests to declare.

References

[1] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

[2] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

[3] Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA)–Thirty years on. *European journal of operational research*, *192*(1), 1-17. https://doi.org/10.1016/j.ejor.2008.01.032

[4] Mergoni, A., Emrouznejad, A., & De Witte, K. (2025). Fifty years of data envelopment analysis. *European journal of operational research*, *326*(3), 389-412. https://doi.org/10.1016/j.ejor.2024.12.049

[5] Banihashemi, S. A., & Khalilzadeh, M. (2021). Time-cost-quality-environmental impact trade-off resource-constrained project scheduling problem with DEA approach. *Engineering, construction and architectural management*, *28*(7), 1979-2004. https://doi.org/10.1108/ECAM-05-2020-0350

[6] Camanho, A. S., Silva, M. C., Piran, F. S., & Lacerda, D. P. (2024). A literature review of economic efficiency assessments using data envelopment analysis (DEA). *European journal of operational research*, *315*(1), 1-18. https://doi.org/10.1016/j.ejor.2023.07.027

[7] Lozano, S., & Villa, G. (2004). Centralized resource allocation using data envelopment analysis. *Journal of productivity analysis*, *22*(1), 143-161. https://doi.org/10.1023/B:PROD.0000034748.22820.33

[8] Asmild, M., Paradi, J. C., & Pastor, J. T. (2009). Centralized resource allocation BCC models. *Omega*, *37*(1), 40-49. https://doi.org/10.1016/j.omega.2006.07.006

[9] Mehdiloozad, M., Sahoo, B. K., & Roshdi, I. (2014). A generalized multiplicative directional distance function for efficiency measurement in DEA. *European journal of operational research*, *232*(3), 679-688. https://doi.org/10.1016/j.ejor.2013.07.042

[10] Fang, H. H., Lee, H. S., Hwang, S. N., & Chung, C. C. (2013). A slacks-based measure of super-efficiency in data envelopment analysis: An alternative approach. *Omega*, *41*(4), 731-734. https://doi.org/10.1016/j.omega.2012.10.004

[11] Seiford, L. M., & Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. *European journal of operational research*, *142*(1), 16-20. https://doi.org/10.1016/S0377-2217(01)00293-4

[12] Färe, R., Grosskopf, S., Lovell, C. K., & Pasurka, C. (1989). Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach. *The review of economics and statistics, 71*(1),90-98. https://doi.org/10.2307/1928055

[13] Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and distance functions. *Journal of economic theory*, *70*(2), 407-419. https://doi.org/10.1006/jeth.1996.0096

[14] Despotis, D. K., & Smirlis, Y. G. (2002). Data envelopment analysis with imprecise data. *European journal of operational research*, *140*(1), 24-36. https://doi.org/10.1016/S0377-2217(01)00200-4

[15] Chen, C., Ren, J., Tang, L., & Liu, H. (2020). Additive integer-valued data envelopment analysis with missing data: A multi-criteria evaluation approach. *PloS one*, *15*(6), e0234247. https://doi.org/10.1371/journal.pone.0234247

[16] Smirlis, Y. G., Maragos, E. K., & Despotis, D. K. (2006). Data envelopment analysis with missing values: An interval DEA approach. *Applied mathematics and computation*, *177*(1), 1-10. https://doi.org/10.1016/j.amc.2005.10.028

[17] Halme, M., Joro, T., Korhonen, P., Salo, S., & Wallenius, J. (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management science*, *45*(1), 103-115. https://doi.org/10.1287/mnsc.45.1.103