buildings

MDPI

Article

Emergency Management Capability Evaluation of Metro Stations Under Earthquake Scenarios from a Resilience Perspective: A Multi-Stage DEA Approach

Linglong Zhou and Heng Yu *

School of Architecture and Civil Engineering & School of Emergency Management, Chengdu University, Chengdu 610106, China; zhoulinglong@cdu.edu.cn
* Correspondence: yuheng@cdu.edu.cn

Abstract

Urban metro systems are highly sensitive to seismic disturbances, and the ability of metro stations to manage emergencies effectively has become an increasingly important component of urban resilience. This study develops a resilience-oriented evaluation framework that conceptualizes emergency management as a sequential managerial process encompassing preparedness, response, and recovery. A multi-dimensional indicator system was constructed based on the four resilience capacities—absorptive, maintaining, recovery, and adaptive—and operationalized through a multi-stage Data Envelopment Analysis (DEA) model. The framework enables both overall efficiency assessment and stage-specific diagnosis of managerial weaknesses. Methodologically, the study demonstrates how resilience theory can be operationalized into a network efficiency structure suitable for process-level diagnosis rather than aggregate scoring. A case study of a representative metro station demonstrates the applicability of the proposed method. The results reveal that while preparedness practices are relatively mature, notable inefficiencies exist in real-time response and post-event recovery due primarily to managerial factors such as communication reliability, personnel coordination, and restoration planning. Improvement simulations confirm that targeted enhancements in these management processes can substantially increase overall emergency efficiency. The findings highlight that seismic resilience is not solely determined by physical infrastructure but is heavily dependent on managerial effectiveness across the emergency cycle. The proposed framework contributes a process-oriented, data-driven tool for evaluating and improving emergency management performance and offers practical guidance for metro operators seeking to strengthen resilience under earthquake scenarios.

Keywords: metro station; earthquake resilience; emergency management; multi-stage DEA; network DEA; management efficiency; preparedness–response–recovery process

Check for updates

Academic Editor: Tarek Zayed

Received: 25 December 2025
Revised: 19 January 2026
Accepted: 24 January 2026
Published: 28 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Research Background

Urban rail transit systems are indispensable components of contemporary metropolitan infrastructure, enabling high-density mobility and supporting the continuity of social and economic activities in large cities [1]. As critical nodes within these systems, metro stations accommodate complex passenger flows, integrate multiple interdependent technical subsystems, and operate in spatially constrained underground environments. From a

resilience and earthquake-engineering perspective, these features increase both exposure and vulnerability, and they intensify dependency across structural, non-structural, and operational functions [2,3].

In earthquake-prone regions, such coupling substantially amplifies seismic risk. Even moderate ground shaking may impair structural integrity, displace critical non-structural components and equipment, interrupt power and communication systems, and hinder safe passenger evacuation [4–6], as conceptually illustrated in Figure 1. Importantly, these disruptions rarely occur in isolation; rather, they can cascade across interdependent subsystems (e.g., ventilation, power supply, signaling, and vertical transportation), so that degradation in one component rapidly propagates into station-level functional loss and constrains emergency response effectiveness [7–9].

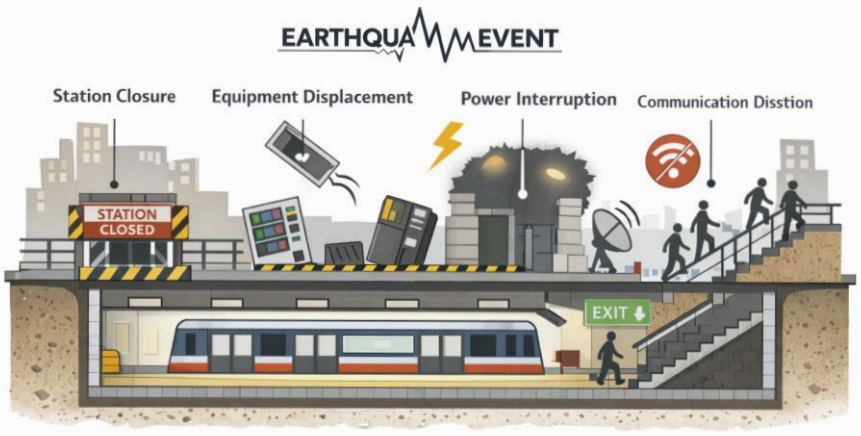

Figure 1. Typical operational disruptions of metro stations under earthquake conditions (Figure Source: generated by authors).

Recent seismic events worldwide further demonstrate that, despite generally stringent seismic design standards, metro stations can remain vulnerable in operational stability and emergency management. For instance, the 2024 earthquake in Taiwan led to temporary metro station closure and visible station-area impacts. Such events reinforce that service disruption, evacuation delays, and emergency resource mobilization challenges are not solely engineering issues but also strongly shaped by management readiness and coordination capacity. In practice, post-event functionality and recovery depend on both structural/non-structural performance and the operational interdependencies that govern system function and restoration [10].

These observations highlight that the overall emergency capability of metro stations is determined not only by physical robustness but also by managerial preparedness, command efficiency, interdepartmental coordination, and recovery planning. Resilience research has repeatedly argued that focusing only on "damage" is insufficient; what matters is the time-dependent trajectory of functionality loss and recovery, and the institutional and organizational means to accelerate restoration [11,12].

Accordingly, traditional engineering-oriented assessments—while necessary—are insufficient to capture the multi-dimensional reality of seismic emergency management in metro stations [13]. As urban rail systems continue to expand in scale and complexity, there is a clear need for comprehensive evaluation frameworks that reflect the dynamic, multi-stage nature of emergency management under seismic conditions, spanning pre-earthquake preparedness, real-time response, and post-earthquake recovery and learning [3,14].

1.2. Research Status and Current Progress

Over the past five years, research on metro stations under earthquake scenarios has grown rapidly, but the literature remains compartmentalized across (i) seismic performance and fragility of underground structures, (ii) passenger evacuation and multi-hazard progression inside stations, (iii) emergency management capacity assessment, and (iv) process-aware evaluation methods capable of diagnosing stage-specific weaknesses. This fragmentation makes it difficult to form an integrated understanding of how "engineering performance → operational disruption → managerial response → recovery trajectory" evolves in a coupled manner during seismic events.

(1) Structural response, fragility, and uncertainty-aware seismic assessment. Recent studies on the seismic performance of underground and metro station structures have moved beyond deterministic response checks toward probabilistic vulnerability and fragility analysis, with increasing attention to soil–structure interaction, parameter uncertainty, and site heterogeneity [15–17]. Research addressing near-fault effects and complex ground-motion characteristics has further improved the understanding of critical demand mechanisms in underground stations [18]. These studies provide an essential foundation for evaluating the physical safety and damage probability of metro stations under earthquake loading. However, from the perspective of emergency management and operational resilience, structural robustness represents only a necessary baseline condition. Even when structural damage remains limited, stations may still experience severe functional disruption due to deficiencies in organization, communication, decision-making, and recovery coordination. This limitation suggests that structural fragility analysis alone is insufficient to explain post-earthquake emergency performance, highlighting the need to incorporate management-oriented and process-based evaluation perspectives.

(2) From structural safety to functionality and resilience at the station scale. Recognizing these limitations, recent research has increasingly shifted from structural safety toward functionality and resilience-oriented assessment at the metro-station scale. Post-earthquake functionality evaluation frameworks have been proposed to quantify service degradation and restoration while explicitly considering interdependencies among subsystems [19]. Related studies integrate infrastructure components and passenger impacts to analyze resilience trajectories under seismic disturbances [20]. In addition, recovery-path-based approaches link functional loss with restoration strategies and recovery speed, enhancing the practical interpretability of resilience outcomes for planning and decision-making [21]. Although these resilience-oriented studies advance the assessment paradigm beyond structural damage states, they generally focus on system-level performance trajectories. The internal management processes that govern preparedness, response, and recovery—and their respective efficiency contributions—are often not explicitly diagnosed.

(3) Evacuation modeling: behavioral realism, guidance strategies, and multi-hazard coupling. Evacuation research for metro stations has evolved from simplified flow-based models toward more behaviorally informed simulations that explicitly examine guidance strategies, control measures, and crowd-distribution mechanisms. Recent studies quantify how different guidance and control schemes influence evacuation effectiveness and passenger movement patterns, providing practical insights for improving station-level evacuation operations [22]. At the same time, there is growing recognition that earthquake response in metro stations rarely involves a single hazard. Instead, it often unfolds under compound or cascading conditions, such as structural damage, smoke or fire, power and communication loss, and congestion, which jointly constrain evacuation and emergency actions. Accordingly, modeling approaches have increasingly emphasized multi-hazard coupling and time-dependent constraints rather than isolated evacuation scenarios [23]. Although these studies significantly advance the understanding of evacuation dynamics

under complex conditions, they primarily focus on movement performance and safety outcomes. The efficiency of emergency management processes—such as command coordination, resource mobilization, and inter-stage linkage between preparedness, response, and recovery—remains largely implicit, limiting their ability to diagnose managerial inefficiencies across the emergency lifecycle.

(4) Digitalization and scenario-based assessment for emergency operations inside metro stations. Another notable trend in recent research is the integration of digitalized station representations, particularly BIM-based models, with simulation tools to support scenario testing, procedural design, and capacity evaluation. BIM-enabled approaches have been widely applied to emergency evacuation safety design in metro stations, improving the consistency among spatial data, operational rules, and simulation environments [24]. Related studies extend beyond evacuation analysis to address broader station-system robustness and resilience management under hazard scenarios, highlighting the potential of digital tools to support integrated emergency decision-making [25]. More generally, BIM-linked simulation systems for large-scale infrastructure have been developed to enhance the credibility, transparency, and usability of evacuation and emergency analyses for decision support [26]. Complementing these digital approaches, indicator-based evacuation capability evaluations provide additional applied evidence on how station configuration and management measures affect evacuation outcomes [27]. While these studies improve scenario realism and operational assessment, they typically remain task- or function-specific, and do not explicitly evaluate the efficiency of emergency management processes across different stages of an earthquake event.

(5) Emergency management capability assessment: progress and limitations. Compared with structural performance and evacuation studies, direct research on the emergency management capability of urban rail transit stations remains relatively less developed, particularly with respect to the coupling among preparedness, response, and recovery stages. Existing studies commonly adopt multi-criteria decision-making approaches, such as AHP–TOPSIS, to construct composite capability indices and benchmarking results for metro lines or stations [28]. Although such approaches are valuable for summarizing overall managerial readiness, they generally treat emergency management as a stage-aggregated and static system. As a result, they offer limited ability to identify where performance bottlenecks emerge along the emergency management chain (preparedness versus response versus recovery), or to explain how inefficiencies propagate across stages under earthquake conditions.

(6) Methodological gap: process-oriented, multi-stage evaluation and the role of network DEA. To diagnose stage-specific inefficiencies and explicitly capture intermediate outputs—such as the transformation from preparedness resources to response effectiveness and ultimately to recovery outcomes—evaluation methods must represent internal process structures and inter-stage dependencies. In operations research, network DEA and its extensions provide a mature methodological foundation for modeling linked stages and intermediate products within complex systems. Recent advances include two-stage network DEA formulations that address production-scale matching issues [29], as well as dynamic network DEA models that incorporate variable time-lag effects, enabling the representation of delayed or carryover impacts in multi-period systems [30]. In addition, systematic reviews of network DEA applications summarize how network structures have been operationalized across diverse domains and highlight methodological options for detailed process diagnosis [31].

Despite this methodological readiness, the application of network or multi-stage DEA to earthquake emergency management in metro stations remains rare. This gap motivates the development of a resilience-informed, process-based framework that (i) explicitly

aligns evaluation stages with preparedness–response–recovery logic, and (ii) quantifies both overall and stage-specific efficiency to pinpoint operational bottlenecks under seismic conditions.

1.3. Research Gap and Objectives

Although existing studies have substantially advanced the understanding of seismic risks in metro systems, several critical challenges remain unresolved. First, most current evaluation approaches conceptualize emergency management as a static system, thereby failing to represent the sequential and interdependent nature of preparedness, response, and recovery processes. This simplification conceals essential causal linkages across emergency stages and limits the ability to diagnose internal inefficiencies and stage-specific bottlenecks. Second, although resilience theory provides a coherent conceptual framework that aligns well with the temporal logic of emergency management, it has not yet been systematically operationalized into measurable and station-specific indicators for metro systems. In particular, the translation of absorptive, maintaining, recovery, and adaptive capacities into a structured and quantifiable assessment framework remains underdeveloped. As a result, resilience is often discussed at a conceptual level, with limited applicability for empirical evaluation and decision support at the station scale. Finally, the methodological tools commonly employed in emergency management evaluation—such as static multi-criteria decision-making techniques—generally lack the capability to model multi-stage operational processes or to quantify efficiency differentials across distinct emergency phases. Consequently, these methods are insufficient for capturing the internal structure and dynamic performance of complex emergency management systems, especially under earthquake-induced cascading effects.

To address these gaps, this study is positioned primarily as a methodological investigation and develops a resilience-oriented evaluation framework for metro station emergency management under earthquake scenarios. The proposed framework constructs a multi-level indicator system grounded in the four core dimensions of seismic resilience and explicitly aligns these indicators with a process-based representation of emergency operations. On this basis, a multi-stage Data Envelopment Analysis (DEA) model is established to characterize the chained input–output relationships linking preparedness, response, and recovery stages.

Rather than aiming to produce statistically generalizable efficiency rankings across multiple stations, the objectives of this study are to (i) demonstrate how resilience theory can be operationalized into a process-based, stage-linked evaluation structure, (ii) illustrate the capability of a multi-stage DEA model to diagnose stage-specific inefficiencies and internal managerial bottlenecks, and (iii) explore targeted improvement directions that enhance emergency management performance along the preparedness–response–recovery chain.

Accordingly, the empirical application to a single metro station is designed as an illustrative case, using scenario-consistent data to validate the methodological logic and diagnostic insight of the proposed framework. The focus is therefore on methodological validity and process-level diagnosis, rather than on operational benchmarking or statistically generalizable efficiency comparison across metro stations.

2. Resilience-Oriented Indicator System

2.1. Indicator System

The evaluation indicator system developed in this study is grounded in resilience thinking, metro-station emergency management mechanisms, and the operational characteristics of earthquake-induced disruptions. Rather than serving as an exhaustive checklist, the system is designed as a theory-informed hierarchical framework intended to represent

how a station absorbs seismic impacts, sustains essential functions during the disturbance, restores degraded performance, and improves through post-event learning. This conceptual stance is consistent with recent resilience assessment research that organizes indicators around capacities spanning absorption/restoration/adaptation and emphasizes the need for measurable attributes that can be validated through structured expert elicitation [32,33].

At the theoretical level, seismic resilience provides the organizing logic for the indicator architecture. Resilience-oriented assessments commonly distinguish (i) the ability to withstand an initial shock with limited performance loss, (ii) the ability to continue delivering critical functions under disruption, (iii) the ability to restore performance efficiently, and (iv) the ability to adapt by institutionalizing lessons learned. Recent work further highlights that these capacities should be operationalized into observable and comparable metrics along a "resilience curve" (impact depth, duration, recovery rate, and improvement gains), which supports the translation of abstract resilience concepts into measurable station-level indicators [32,33].

Within this resilience-based structure, the criterion layer is mapped to core mechanisms that shape emergency management outcomes across the disaster lifecycle. Absorptive capacity reflects the station's pre-event robustness and readiness, including structural and equipment-related protective measures, redundancy and resource sufficiency, and personnel preparedness. The importance of explicitly capturing preparedness elements (organization, monitoring/early warning, emergency support, and readiness-related arrangements) has been emphasized in systematic syntheses of emergency preparedness practice in rail-transit contexts, which argue that preparedness quality strongly conditions downstream response and recovery effectiveness [34].

Maintaining capacity focuses on the station's ability to sustain essential functions during the disruptive period, emphasizing timely command actions, communication stability, staff mobilization, and real-time coordination. In underground and station environments, the operational consequence of these factors is often manifested through passenger management performance and evacuation controllability; accordingly, evacuation-capability research has increasingly incorporated management- and organization-related indicators (in addition to human/equipment/environment factors) when constructing station evaluation index systems [35].

Recovery capacity represents the station's ability to restore degraded functions after the event, including clearance and repair efficiency, restoration sequencing of critical subsystems, and service resumption. Recent metro resilience studies demonstrate that recovery strategies and the prioritization of station/component restoration can materially affect system-level resilience outcomes, underscoring that recovery indicators should capture both speed and strategy effectiveness rather than only "time-to-restore" [36,37].

Adaptive capacity captures the longer-term learning and improvement mechanisms that follow emergencies, such as revising emergency plans, improving training and exercises, upgrading technologies, and strengthening inter-organizational coordination. Empirical analyses of rail-transit emergency response plans show that plan quality, internal consistency, and completeness vary substantially across jurisdictions, implying that "learning and plan iteration" can be operationalized as observable adaptive indicators rather than treated as a purely qualitative afterthought [38].

Accordingly, the indicator system is specified as a hierarchical structure with four resilience dimensions at the criterion layer, each populated by secondary indicators selected for theoretical relevance, empirical measurability, and operational significance. The selection and definition of indicators follow three principles. First, indicators must align with the functional logic of metro-station emergency management (e.g., resource deployment, coordination, information reliability, and service continuity), so that each indicator corre-

sponds to an actionable responsibility. Second, preference is given to indicators that can be quantified using operational records, field inspections, or structured expert judgement, and this approach is widely adopted in resilience assessment frameworks seeking comparability across systems [32].

Third, indicators should reflect both deterministic attributes (e.g., redundancy, protective configurations) and dynamic process characteristics (e.g., timeliness, restoration trajectories, learning effectiveness), ensuring the system can support process-oriented diagnostic evaluation rather than static scoring alone.

Finally, because this study proceeds to a multi-stage (network) efficiency diagnosis, the indicator design also considers the need to represent chained input–intermediate–output relationships across emergency stages. Recent developments in network DEA and dynamic/network DEA modeling emphasize that internal structure and intermediate products should be explicitly represented to avoid masking stage-specific inefficiencies [39–41], and that multi-indicator systems can be constructed to evaluate performance along emergency lifecycles using DEA-type models [42].

Table 1 summarizes the secondary indicators and their definitions under the absorptive capacity dimension. These indicators capture structural robustness, equipment anchoring quality, redundancy of essential systems, adequacy of emergency supplies, and the preparedness of personnel through training and drills. They represent fundamental conditions influencing the system's ability to absorb the initial impact of seismic events.

Table 1. Secondary Indicators under Absorptive Capacity (A1).

Indicator	Symbol	Suggested Quantification/Standard Reference	Unit	DEA Role
Structural robustness level	$A1_1$	Compliance with seismic design standards; evaluation of structural vulnerability	score	Input (Preparedness stage)
Equipment anchorage integrity	$A1_2$	Proportion of critical equipment meeting seismic anchoring requirements	%	Input (Preparedness stage)
Emergency resource adequacy	$A1_3$	Ratio of available emergency materials to standard requirements	%	Input (Preparedness stage)
Staff training completeness	$A1_4$	Proportion of trained staff relative to operational personnel	%	Input (Preparedness stage)
Earthquake plan maturity	$A1_5$	Expert evaluation of plan completeness, clarity, and operability	score	Intermediate output (Preparedness → Response)

Table 2 presents the indicators associated with maintaining capacity. These indicators assess the station's real-time operational capabilities during an earthquake, including command responsiveness, communication stability, evacuation coordination effectiveness, and the robustness of key systems. Maintaining capacity is essential for preventing cascading failures and ensuring orderly evacuation.

Table 2. Secondary Indicators under Maintaining Capacity (A2).

Indicator	Symbol	Suggested Quantification	Unit	DEA Role
Command response time	$A2_1$	Time between seismic alert and operational command issuance	s	Output (Response stage)
Communication system availability	$A2_2$	Proportion of communication channels functioning during the event	%	Output (Response stage)
Evacuation coordination effectiveness	$A2_3$	Expert scoring of coordination quality among staff	score	Output (Response stage)
Operational continuity of key subsystems	$A2_4$	Percentage of critical functions maintained during the earthquake	%	Output (Response stage)
Emergency staffing adequacy	$A2_5$	Availability ratio of on-duty emergency personnel	%	Intermediate output (Response → Recovery)

The indicators under recovery capacity, shown in Table 3, capture the efficiency of post-earthquake restoration activities. This dimension reflects the system's ability to reduce

downtime, limit operational losses, and restore normal operation within acceptable timeframes. Recovery capacity thus serves as a direct measure of the practical implementation of resilience.

Table 3. Secondary Indicators under Recovery Capacity (A3).

Indicator	Symbol	Suggested Quantification/Standard Reference	Unit	DEA Role
Time to restore critical equipment	$A3_1$	Duration from earthquake to functional restoration	h	Output (Recovery stage)
Service recovery level	$A3_2$	Ratio of restored operational capacity to normal capacity	%	Output (Recovery stage)
Environmental clearance efficiency	$A3_3$	Time required to remove debris and secure passenger areas	min	Output (Recovery stage)
Resource deployment timeliness	$A3_4$	Comparison of actual deployment to the planned schedule	score	Output (Recovery stage)
Passenger service restoration time	$A3_5$	Time required to resume passenger services	h	Output (Recovery stage)

Although several indicators in the proposed system relate to resource management, they are intentionally defined to capture distinct stages and mechanisms of the emergency management process, rather than overlapping information. Specifically, emergency resource adequacy reflects the baseline sufficiency of resources prior to an earthquake (preparedness stage), resource deployment timeliness captures the speed and effectiveness of mobilizing available resources during post-event response and recovery, and improvement in resource allocation mechanisms represents longer-term adaptive enhancement of allocation strategies based on post-event learning. These indicators therefore differ in temporal focus (pre-event, during-event, and post-event adaptation) and functional role within the preparedness–response–recovery chain. By assigning them to different stages in the multi-stage DEA framework, the model avoids excessive redundancy while enabling diagnosis of how resource-related inefficiencies emerge and propagate across stages.

Adaptive capacity indicators, presented in Table 4, examine how metro stations incorporate lessons learned from seismic events into long-term management improvements. This dimension is crucial for enhancing resilience over time, ensuring that emergency management evolves in response to emerging risks, operational feedback, and technological innovations.

Table 4. Secondary Indicators under Adaptive Capacity (A4).

Indicator	Symbol	Suggested Quantification/Standard Reference	Unit	DEA Role
Emergency plan revision frequency	$A4_1$	Number of updates per year	times/year	Output (Recovery/Adaptation stage)
Effectiveness of post-event review	$A4_2$	Expert assessment of the depth and applicability of after-action reviews	score	Output (Recovery/Adaptation stage)
Improvement in resource allocation mechanisms	$A4_3$	Comparison of pre- and post-event resource deployment efficiency	%	Output (Recovery/Adaptation stage)
Integration of new technologies	$A4_4$	The degree to which intelligent systems enhance emergency management	score	Output (Recovery/Adaptation stage)
Organizational learning effectiveness	$A4_5$	Assessment of institutional learning processes and staff adoption	score	Output (Recovery/Adaptation stage)

From an operational perspective, the indicators in the proposed hierarchical system differ in their primary data sources and measurement requirements. A substantial proportion of the indicators—particularly those related to equipment condition, system availability, response timing, staffing levels, and recovery duration—can be obtained from routine operational records of a typical metro agency, such as maintenance logs, emergency plans, incident reports, dispatch records, and system monitoring data. These indicators reflect objective operational performance and are commonly documented during drills or real emergency events.

Other indicators, especially those associated with plan maturity, coordination effectiveness, post-event review quality, and organizational learning, rely primarily on structured expert judgement. Such indicators capture qualitative aspects of emergency management that are difficult to observe directly through operational records alone and can be elicited through standardized expert scoring procedures, expert workshops, or Delphi-style evaluations, which are widely adopted in emergency management and resilience assessment practice.

By organizing the indicator system in this manner, the proposed framework captures both static preparedness attributes and dynamic operational characteristics of metro station emergency management. Each indicator is explicitly aligned with a specific phase of the resilience cycle, ensuring compatibility with the multi-stage DEA model introduced later. The explicit distinction between indicators derived from operational records and those based on expert judgement enhances transparency and practical implementability. Overall, the coherence between theoretical foundations, operational mechanisms, and data acquisition pathways allows the indicator system to function not only as a conceptually robust structure but also as a realistically applicable analytical tool for evaluating seismic emergency management efficiency in metro stations.

2.2. Indicator Value Acquisition

2.2.1. Absorptive Capacity Indicators ($A1$)

To support rigorous quantitative assessment, each secondary indicator defined in Section 2.1 is translated into a clearly specified numerical value. This involves identifying appropriate data sources, defining measurement procedures, and formulating explicit calculation methods. For each indicator below, a short conceptual description is followed by the computational formula, a detailed explanation of variables and units, and a brief statement of the indicator's role in the resilience framework.

(1) Structural robustness level ($A1_1$). Unit: dimensionless score (0–1). Structural robustness is assessed by evaluating the extent to which the main structural components of the station comply with seismic design standards and exhibit low vulnerability in service. A structured checklist is applied to elements such as slabs, beams, columns, load-bearing walls, seismic joints, and waterproofing details. Each item is scored for seismic compliance and in-service condition; the station's overall robustness level is then obtained as a weighted average of item scores.

$$A1_1 = \frac{\sum_{k=1}^{m} \omega_k S_k}{\sum_{k=1}^{m} \omega_k} \tag{1}$$

In which, S_k (dimensionless, score 0–1): compliance and condition score for structural item k. ω_k (dimensionless): importance weight of structural item k. m (dimensionless, count): total number of structural items evaluated. This indicator reflects the inherent structural capacity of the station to absorb seismic forces without severe damage.

(2) Equipment anchorage integrity ($A1_2$). Unit: percentage (%). Equipment anchorage integrity is assessed by determining the proportion of critical equipment that satisfies seismic anchoring requirements. A list of key items (e.g., switchgear, transformers, electrical cabinets, ventilation units, communication racks) is compiled, and on-site inspections verify whether each item is anchored according to relevant seismic standards. This indicator reflects the physical stability of critical equipment and its ability to withstand shaking without displacement or overturning.

$$A1_2 = \frac{N_{\text{ok}}}{N_{\text{tot}}} \tag{2}$$

In which, N_{ok} (count: items): number of equipment items whose anchorage meets seismic requirements. N_{tot} (count: items): total number of critical equipment items inspected.

(3) Emergency resource adequacy ($A1_3$). Unit: percentage (%). Emergency resource adequacy is assessed by comparing available emergency materials with standard requirements specified in the station's emergency plan. For each resource category (e.g., first-aid kits, fire extinguishers, emergency lights, protective equipment, backup communication devices), an adequacy ratio is calculated and then aggregated into a station-level value.

$$\alpha_i = \min\left(\frac{Q_i^{\text{avail}}}{Q_i^{\text{req}}},\ 1\right) \tag{3}$$

$$A1_3 = \frac{\sum_{i=1}^{n} \omega_i \alpha_i}{\sum_{i=1}^{n} \omega_i} \tag{4}$$

In which, Q_i^{avail} (various units: pieces/sets/kg, depending on resource type): available quantity of resource category i. Q_i^{req} (same units as Q_i^{avail}): required quantity for resource category i according to the emergency plan. α_i (dimensionless, 0–1): adequacy ratio of resource category i. ω_i (dimensionless): importance weight of resource category i. n (count): number of resource categories included in the calculation.

The weighting coefficients (ω_i, ω_k) used to aggregate indicator components reflect their relative importance within a given capacity dimension. In practical applications, these weights can be determined through structured expert judgement (e.g., expert scoring, Delphi consultation, or AHP) based on operational experience and management priorities. Alternatively, when sufficient data are available, objective approaches such as equal weighting may be adopted. In this study, scenario-consistent weights are used for illustrative purposes.

(4) Staff training completeness ($A1_4$). Unit: percentage (%). Personnel readiness is assessed by determining the proportion of staff who have completed mandatory training modules within the prescribed time frame. Training modules typically include evacuation guidance, hazard identification, communication protocols, and seismic equipment operation.

$$A1_4 = \frac{N_{\text{train}}}{N_{\text{staff}}} \tag{5}$$

In which, N_{train} (count: persons): number of personnel who have completed all required earthquake-related emergency training within the reference period. N_{staff} (count: persons): total number of operational and emergency-related staff assigned to the station.

(5) Earthquake plan maturity ($A1_5$). Unit: dimensionless score (0–1). Earthquake plan maturity is assessed through expert evaluation of the station's emergency plan in terms of its completeness, clarity, operability, and consistency with higher-level plans. Experts score several dimensions of plan quality using a standardized rating scale; the scores are then normalized and averaged.

$$A1_5 = \frac{1}{J}\sum_{j=1}^{J} \frac{E_j}{E_{max}} \tag{6}$$

In which, E_j (dimensionless, score): expert score for plan quality dimension j (e.g., clarity of roles, communication procedures, scenario coverage). E_{max} (dimensionless, score): maximum possible score on the rating scale for each dimension. J (count): number of plan quality dimensions evaluated.

2.2.2. Maintaining Capacity Indicators ($A2$)

(1) Command response time ($A2_1$). Seconds (raw time); dimensionless score (0–1) after conversion. Command response time is assessed by measuring the elapsed time between

the recorded seismic alert and the issuance of the first formal operational command (e.g., initiating evacuation, switching to emergency mode, suspending services). To make the indicator compatible with benefit-type outputs in DEA, the raw time is converted into a normalized score, where shorter times correspond to higher values.

$$A2_1 = 1 - \frac{T_{\text{cmd}} - T_{min}}{T_{max} - T_{min}} \tag{7}$$

In which, T_{cmd} (seconds, s): observed command response time at the station. T_{min} (seconds, s): reference minimum response time representing best practice. T_{max} (seconds, s): reference maximum response time representing the worst acceptable performance.

For time-based indicators normalized into 0–1 scores, T_{min} and T_{max} define the reference performance bounds. In practical applications, T_{min} can be set as a best-practice or target value based on historical operational records, internal standards, or emergency response guidelines, while T_{max} represents a worst acceptable upper bound, derived from emergency plans, regulatory thresholds, or observed extremes in past events or drills.

In this study, scenario-consistent reference values are adopted to illustrate the normalization procedure. Since the normalization functions are monotonic, moderate variations in T_{min} and T_{max} affect score scaling but do not alter the qualitative diagnosis of stage-specific inefficiencies.

The resulting efficiency scores for time-based indicators (e.g., $A2_1$, $A3_1$, $A3_5$) should be interpreted relative to the selected reference bounds T_{min} and T_{max}. When observed values are close to T_{min}, the corresponding scores indicate near-best-practice performance consistent with established emergency response targets. Conversely, values approaching T_{max} reflect delayed response or recovery beyond acceptable operational thresholds. Under the adopted reference bounds, the obtained scores therefore provide a normalized and operationally interpretable measure of time efficiency, enabling meaningful comparison across stages within the same station rather than absolute performance benchmarking.

(2) Communication system availability ($A2_2$). Unit: percentage (%). Communication system availability is assessed by determining the proportion of communication channels that remain functional during the seismic event. Channels may include fixed telephones, radios, digital dispatch terminals, and emergency intercoms specified in the emergency plan.

$$A2_2 = \frac{C_{\text{ok}}}{C_{\text{tot}}} \tag{8}$$

In which, C_{ok} (count: channels): number of communication channels operating normally during the earthquake or drill. C_{tot} (count: channels): total number of communication channels required for emergency management.

(3) Evacuation coordination effectiveness ($A2_3$). Unit: dimensionless score (0–1). Evacuation coordination effectiveness is assessed by expert observers during drills or scenario-based exercises. Experts evaluate how well staff coordinate to guide passengers, manage congestion, and handle unexpected situations, and assign scores using a structured rubric.

$$A2_3 = \frac{1}{N}\sum_{i=1}^{N} \frac{S_i}{S_{max}} \tag{9}$$

In which, S_i (dimensionless, score): evacuation coordination score assigned by expert i. S_{max} (dimensionless, score): maximum possible score on the rubric. N (count: persons): number of experts participating in the evaluation.

(4) Operational continuity of key subsystems ($A2_4$). Unit: percentage (%) for each γ_k; overall indicator expressed as a percentage or dimensionless average. Operational continuity of key subsystems is assessed by measuring the proportion of time that critical

subsystems remain functional during the seismic event. These subsystems typically include traction power, auxiliary power, ventilation, lighting, ticketing, and signaling.

$$\gamma_k = \frac{T_k^{\text{on}}}{T^{\text{event}}}, \tag{10}$$

$$A2_4 = \frac{1}{K}\sum_{k=1}^{K} \gamma_k \tag{11}$$

In which, T_k^{on} (seconds or minutes): duration for which subsystem k remains operational during the critical period of the earthquake. T^{event} (seconds or minutes): total duration of the critical period considered (e.g., from onset of shaking to completion of immediate response). γ_k (dimensionless): continuity ratio for subsystem k. K (count): number of key subsystems included in the assessment.

(5) Emergency staffing adequacy ($A2_5$). Unit: percentage (%). Emergency staffing adequacy is assessed by comparing the number of staff actually available during the earthquake with the number specified in the emergency plan as necessary to perform key response roles. Availability includes staff on duty and those who can be rapidly mobilized within the required response timeframe.

$$A2_5 = \frac{N_{\text{avail}}}{N_{\text{req}}} \tag{12}$$

In which, N_{avail} (count: persons): number of staff available and able to assume emergency roles during the event. N_{req} (count: persons): number of staff required for full coverage of emergency roles according to the plan.

2.2.3. Recovery Capacity Indicators ($A3$)

(1) Time to restore critical equipment ($A3_1$). Unit: hours (raw time); dimensionless score (0–1) after conversion. Time to restore critical equipment is assessed by measuring the duration between the end of the seismic event (or the issuance of an "all clear" for structural safety) and the moment when each critical device is confirmed to be operational again. Critical equipment includes power supply units, ventilation systems, lighting circuits, and key control devices. A representative station-level restoration time is obtained (e.g., the maximum or a weighted average across critical devices) and then normalized.

$$A3_1 = 1 - \frac{T_{\text{rest}} - T_{min}}{T_{max} - T_{min}} \tag{13}$$

In which, T_{rest} (hours, h): observed time required for restoration of critical equipment at the station. T_{min} (hours, h): reference minimum restoration time representing the best achievable performance. T_{max} (hours, h): reference maximum restoration time representing the worst acceptable performance.

(2) Service recovery level ($A3_2$). Unit: percentage (%). Service recovery level is assessed by comparing the operational capacity achieved after the earthquake with the normal capacity under routine conditions. Capacity is typically measured as train frequency (trains per hour) or passenger throughput (passengers per hour) at a specified reference time in the recovery phase.

$$A3_2 = \frac{Q_{\text{rec}}}{Q_{\text{norm}}} \tag{14}$$

In which, Q_{rec} (trains/h or passengers/h): recovered service capacity at the defined recovery time point. Q_{norm} (trains/h or passengers/h): normal scheduled or design service capacity.

(3) Environmental clearance efficiency ($A3_3$). Unit: minutes (raw time); dimensionless score (0–1) after conversion. Environmental clearance efficiency is assessed by determining how long it takes to remove debris, secure loose objects, and complete safety inspections in passenger and staff areas after the seismic event. The raw clearance time is then transformed into a performance score so that faster clearance corresponds to higher resilience.

$$A3_3 = 1 - \frac{T_{\text{clear}} - T_{min}}{T_{max} - T_{min}} \tag{15}$$

In which, T_{clear} (minutes, min): observed time taken to complete environmental clearance and safety checks. T_{min} (minutes, min): reference minimum clearance time representing best practice. T_{max} (minutes, min): reference maximum clearance time representing the upper bound of acceptable performance.

(4) Resource deployment timeliness ($A3_4$). Unit: minutes (raw time); dimensionless score (0–1) after conversion. Resource deployment timeliness is assessed by comparing the actual timing of mobilizing key recovery resources (such as maintenance crews, inspection teams, or external support units) with the planned schedule defined in the emergency plan. The closer the actual deployment is to the planned target, the higher the timeliness score. If time data are available, the indicator can be expressed as:

$$A3_4 = 1 - \frac{T_{\text{act}} - T_{\text{plan}}}{T_{max} - T_{\text{plan}}} \tag{16}$$

In which, T_{act} (minutes, min): actual deployment or arrival time of the specified resource. T_{plan} (minutes, min): planned deployment time according to the emergency plan. T_{max} (minutes, min): maximum acceptable deployment time used as an upper bound.

(5) Passenger service restoration time ($A3_5$). Unit: hours (raw time); dimensionless score (0–1) after conversion. Passenger service restoration time is assessed by measuring the time interval between the end of the seismic emergency and the resumption of passenger service at the station (either in provisional or normal form). As with other time-based indicators, the raw restoration time is converted into a normalized performance score.

$$A3_5 = 1 - \frac{T_{\text{psr}} - T_{min}}{T_{max} - T_{min}} \tag{17}$$

In which, T_{psr} (hours, h): observed time required to resume passenger services after the earthquake. T_{min} (hours, h): reference minimum service restoration time representing best practice. T_{max} (hours, h): reference maximum service restoration time representing the worst acceptable performance.

2.2.4. Adaptive Capacity Indicators ($A4$)

(1) Emergency plan revision frequency ($A4_1$). Unit: revisions per year (times/year). Emergency plan revision frequency is assessed by counting how many times the station's earthquake-related emergency plan has undergone substantive updates within a given observation period. Substantive updates include significant changes in procedures, responsibilities, communication protocols, or resource allocation rules.

$$A4_1 = \frac{N_{\text{rev}}}{Y} \tag{18}$$

In which, N_{rev} (count: revisions): number of substantive plan revisions recorded. Y (years): length of the observation period.

(2) Effectiveness of post-event review ($A4_2$). Unit: dimensionless score (0–1). The effectiveness of post-event review is assessed by expert evaluation of the quality and depth

of after-action reviews conducted following drills or real earthquakes. A structured rubric considers dimensions such as thoroughness of incident reconstruction, identification of root causes, specificity of recommended corrective actions, and follow-up implementation.

$$A4_2 = \frac{1}{M}\sum_{i=1}^{M}\frac{A_i}{A_{max}} \tag{19}$$

In which, A_i (dimensionless, score): expert assessment score for review dimension i. A_{max} (dimensionless, score): maximum possible score for each dimension. M (count): number of dimensions used to evaluate review effectiveness.

(3) Improvement in resource allocation mechanisms ($A4_3$). Unit: percentage (%). Improvement in resource allocation mechanisms is assessed by comparing selected performance metrics (such as average response time or average equipment restoration time) before and after the implementation of new allocation strategies. The indicator expresses the relative improvement as a percentage.

$$A4_3 = \frac{T_{\text{before}} - T_{\text{after}}}{T_{\text{before}}} \tag{20}$$

In which, T_{before} (minutes or hours): average value of a chosen performance metric (e.g., response time) before resource allocation improvements are implemented. T_{after} (minutes or hours): average value of the same performance metric after improvements.

(4) Integration of new technologies ($A4_4$). Unit: dimensionless score (0–1). Integration of new technologies is assessed by examining the extent to which intelligent systems—such as seismic monitoring interfaces, real-time passenger flow analytics, digital command platforms, and automated alert systems—have been adopted and embedded into emergency management procedures. Each technology type is rated for its level of adoption and integration, and a weighted average is used to derive the station-level indicator.

$$A4_4 = \frac{\sum_{k=1}^{K}\omega_k U_k}{\sum_{k=1}^{K}\omega_k} \tag{21}$$

In which, U_k (dimensionless, score): adoption and integration score for technology type k. ω_k (dimensionless): importance weight assigned to technology type k. K (count): number of technology types considered.

(5) Organizational learning effectiveness ($A4_5$). Unit: dimensionless score (0–1). Organizational learning effectiveness is assessed by evaluating the degree to which lessons identified from drills and real events are institutionalized in the form of updated procedures, revised training content, and observable changes in staff behavior. Experts rate several learning-related dimensions, such as the existence of formal feedback loops, staff participation in learning activities, and evidence of implemented improvements.

$$A4_5 = \frac{1}{L}\sum_{j=1}^{L}\frac{O_j}{O_{max}} \tag{22}$$

In which, O_j (dimensionless, score): expert score for learning dimension j (e.g., feedback mechanisms, training updates, behavior change). O_{max} (dimensionless, score): maximum possible score on the rating scale. L (count): number of learning dimensions evaluated.

3. Multi-Stage DEA Model for Resilience-Oriented Evaluation

3.1. Overall Evaluation Framework

After defining the resilience-oriented indicator system, the next step is to develop a quantitative method capable of integrating the indicators and revealing the efficiency of

emergency management across different stages of an earthquake event. In this study, a multi-stage Data Envelopment Analysis (DEA) model is adopted to evaluate metro station emergency management capability from a resilience perspective. Although DEA is typically applied to multiple DMUs, it has also been widely adopted as a process-diagnostic tool in methodological studies, where virtual reference frontiers are constructed to explore internal inefficiencies rather than to rank entities. The core idea of this study is to represent the emergency management process as a chained series of production stages. Metro stations are treated as decision-making units (DMUs) that transform preparedness inputs into intermediate capabilities before an earthquake, convert those capabilities into effective response performance during the event, and finally transform response performance into recovery outcomes after the event. This structure aligns with the resilience dimensions and with the practical operations of metro emergency management.

In the proposed framework, the emergency process is divided into three stages. The first stage corresponds to pre-earthquake preparedness, focusing on absorptive capacity. This stage uses resource-related indicators such as staff, training, equipment anchorage, and emergency supplies as inputs, and generates preparedness-related intermediate outputs such as readiness scores and system robustness indices. The second stage corresponds to emergency response during the earthquake, emphasizing maintaining capacity. It takes the intermediate outputs from the first stage as inputs, combined with certain real-time resource indicators, and produces response-related outputs such as command timeliness, communication availability, and evacuation coordination performance. The third stage corresponds to post-earthquake recovery, associated mainly with recovery and adaptive capacities. This stage uses response performance indicators as inputs and yields final outputs that reflect the speed and quality of service restoration, recovery of critical equipment, and improvements in management mechanisms.

By modeling the emergency process as a series of stages with intermediate products, the multi-stage DEA framework captures the internal structure of the emergency management system rather than treating it as a black box. The overall efficiency of each metro station is evaluated in terms of its ability to convert initial preparedness inputs into desirable final outcomes through the chain of intermediate stages. At the same time, stage-specific efficiency measures provide insight into which part of the emergency chain is most responsible for overall inefficiency. The evaluation procedure adopted in this study follows a logical sequence. Metro stations are first selected as DMUs, and the indicators in the resilience-oriented system are assigned to the three stages according to their functional roles. Input, intermediate, and output matrices are then constructed for all DMUs. A suitable orientation and returns-to-scale assumption are specified based on the evaluation objectives and data characteristics. The multi-stage DEA model is solved to obtain overall and stage-wise efficiency scores. Finally, these scores are interpreted and classified to identify efficient and inefficient DMUs and to diagnose bottlenecks in the emergency management process.

When applied in a setting with a limited number of decision-making units, the proposed multi-stage DEA model is intended to function as a process-diagnostic and benchmarking tool rather than a statistical efficiency-ranking instrument. In this context, the efficiency frontier represents a reference-efficient configuration derived from best-practice combinations within the modeled emergency management process. The resulting efficiency scores are therefore interpreted as indicators of internal performance gaps and improvement potential along different emergency stages, rather than as comparative rankings across stations.

3.2. Model Formulation

In general, the proposed framework accommodates n metro stations as decision-making units (DMUs), each undergoing three sequential stages of emergency management: preparedness, response, and recovery. For the j-th DMU $(j = 1, 2, \ldots, n)$, the first stage employs m_1 inputs denoted by $x_{ij}^{(1)}$ and generates q_1 intermediate outputs denoted by $z_{rj}^{(1)}$. The second stage uses m_2 inputs, which may include both intermediate outputs from the first stage and additional stage-specific inputs, to produce q_2 intermediate outputs denoted by $z_{sj}^{(2)}$. The third stage then transforms m_3 inputs—incorporating intermediate measures from the second stage—into q_3 final outputs y_{kj}.

In line with the objective of assessing how effectively metro stations convert available emergency-management resources into resilience-related outcomes, an input-oriented DEA model with variable returns to scale (VRS) is adopted. The input-oriented specification is appropriate because metro station operators generally have greater managerial control over organizational and resource-related inputs—such as staffing, training, equipment readiness, and emergency procedures—than over external factors such as earthquake intensity or passenger demand during an event. From an emergency-management perspective, the evaluation therefore focuses on how efficiently controllable inputs are transformed into desirable response and recovery outcomes. The VRS assumption reflects the substantial heterogeneity among metro stations in terms of size, functional role, and operational complexity; large interchange stations and smaller local stations do not necessarily exhibit proportional input–output relationships, particularly under emergency conditions. Allowing for variable returns to scale thus provides a more realistic representation of scale heterogeneity than a constant returns to scale (CRS) specification.

For the overall system, the DEA framework constructs a virtual composite DMU as a linear combination of observed DMUs. Efficiency is evaluated by comparing each DMU with this composite benchmark, such that the benchmark uses no more input while producing at least the same level of outputs at each stage. The resulting efficiency score represents the maximum proportional reduction in inputs that the evaluated DMU could achieve while remaining feasible within the production possibility set.

For a given DMU_o, the overall input-oriented efficiency can be obtained by solving the following linear programming problem, which takes a network structure into account:

$$min_{\theta, \lambda_j^{(s)}, s^-, s^+}^{\theta} \tag{23}$$

Subject to

$$\sum_{j=1}^{n} \lambda_j^{(1)} x_{ij}^{(1)} + s_i^{-(1)} = \theta x_{io}^{(1)}, \forall i \tag{24}$$

$$\sum_{j=1}^{n} \lambda_j^{(1)} z_{rj}^{(1)} - s_r^{+(1)} = z_{ro}^{(1)}, \forall r \tag{25}$$

$$\sum_{j=1}^{n} \lambda_j^{(2)} x_{ij}^{(2)} + s_i^{-(2)} = x_{io}^{(2)}, \forall i \tag{26}$$

$$\sum_{j=1}^{n} \lambda_j^{(2)} z_{rj}^{(2)} - s_r^{+(2)} = z_{so}^{(2)}, \forall s \tag{27}$$

$$\sum_{j=1}^{n} \lambda_j^{(3)} x_{ij}^{(3)} + s_i^{-(3)} = x_{io}^{(3)}, \forall i \tag{28}$$

$$\sum_{j=1}^{n} \lambda_j^{(3)} y_{ki} - s_k^{+(3)} = y_{ko}, \forall k \tag{29}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} = 1, s = 1, 2, 3 \tag{30}$$

$$\lambda_j^{(s)} \geq 0, s_i^{-(s)} \geq 0, s_r^{+(s)} \geq 0 \tag{31}$$

In these equations, θ is the overall efficiency score of DMU_o; $\lambda_j^{(s)}$ are the intensity variables associated with the j-th DMU in stage s, s^-, and s^+; represent input excesses and output shortfalls (slacks), respectively; and the convexity constraints ensure variable returns to scale in each stage. The constraints describe the feasible set for the three stages, linked through intermediate outputs. When $\theta = 1$ and all slacks are zero, the DMU lies on the efficient frontier; when $\theta < 1$, the DMU is inefficient and could reduce inputs proportionally by a factor of $1 - \theta$ while maintaining its current level of outputs.

To obtain stage-specific efficiency scores, the network structure can be decomposed so that each stage is evaluated conditionally on the intermediate products coming from adjacent stages. For example, the efficiency of the preparedness stage can be evaluated by treating its inputs and intermediate outputs as the sole input–output pair in a conventional DEA model:

$$min_{\theta^{(1)},\lambda_j,s^-,s^+}^{\theta^{(1)}} \tag{32}$$

Subject to:

$$\sum_{j=1}^{n} \lambda_j x_{ij}^{(1)} + s_i^- = \theta^{(1)} x_{io}^{(1)}, \forall i \tag{33}$$

$$\sum_{j=1}^{n} \lambda_j z_{rj}^{(1)} - s_r^+ = z_{ro}^{(1)}, \forall r \tag{34}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, s_i^- \geq 0, s_r^+ \geq 0 \tag{35}$$

Analogous formulations can be derived for the response and recovery stages, with their own inputs and outputs. This decomposition allows the analyst to compute stage efficiencies $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}$, and compare them with the overall efficiency θ. The comparison reveals whether inefficiencies originate mainly from inadequate preparedness, suboptimal response, or delayed recovery.

In practical implementation, all continuous indicators are used directly in the model after appropriate normalization, if necessary, whereas qualitative indicators quantified by expert scores are treated as numeric variables. To avoid scale distortions, variables are standardized or transformed where needed, but the relative relationships among DMUs remain unchanged, preserving the validity of the DEA results.

In the illustrative case study, the empirical analysis focuses on a single metro station as the physical entity under evaluation. In this context, the role of DEA is not to construct a conventional empirical efficiency frontier across multiple stations, but to exploit the internal multi-stage structure of the emergency management process to enable process-level diagnosis.

Specifically, the efficiency frontier is implicitly defined through the network DEA formulation, where each stage is constrained by its own input–output relationships and by the linking conditions imposed by intermediate products. A virtual composite DMU is constructed through the linear programming formulation, representing an internally consistent benchmark within the feasible production possibility set. The resulting efficiency scores therefore indicate the extent to which inputs at each stage could be proportionally reduced while maintaining observed intermediate and final outputs, rather than relative performance ranking across stations.

This setting is appropriate for demonstrating the methodological validity and diagnostic capability of the proposed framework, while multi-station datasets are required for conventional cross-unit efficiency comparison.

3.3. Data Processing and Mapping of Indicators to Stages

For the multi-stage DEA model to be operational, the resilience-oriented indicators must be mapped explicitly to the inputs and outputs of each stage. This mapping is guided by the functional roles of the indicators in the emergency process.

Indicators related to structural robustness, equipment anchorage, emergency supplies, and staff training are assigned as inputs to the preparedness stage. These variables represent resource investments and baseline capacities prior to an earthquake. The outputs of this stage are aggregated preparedness indices, such as composite scores of plan maturity or readiness levels, which are constructed by integrating several underlying indicators through weighted aggregation or expert assessment. These aggregated indicators serve as intermediate products that enter the response stage as inputs.

In the response stage, indicators such as command response time, communication system availability, and evacuation coordination effectiveness are treated as outputs. Emergency staffing adequacy is treated as an output of the response stage and an intermediate input to the recovery stage, reflecting the effectiveness of personnel mobilization during emergency response and its direct influence on subsequent recovery operations. The outputs of the response stage summarize the quality and timeliness of immediate emergency actions, which in turn shape the initial conditions under which recovery begins.

The recovery stage uses response performance indicators as inputs, together with any additional resources mobilized during the recovery period, and generates outputs that reflect restoration outcomes. These outputs include the time to restore critical equipment, the level of service recovery, and the speed of passenger service resumption. Adaptive capacity is modeled as a delayed-output of the recovery stage, consistent with the logic of post-event institutional learning. Selected adaptive capacity indicators—such as the effectiveness of post-event reviews or the integration of new technologies—can either be incorporated into the output set of the recovery stage or treated as separate evaluation variables linked to long-term resilience.

Prior to model computation, all indicator data are checked for completeness and consistency. Outliers, if any, are examined to confirm whether they reflect measurement errors or genuine performance differences. Where necessary, monotonic transformations are applied to ensure that all variables have a "the smaller, the better" or "the larger, the better" direction consistent with the DEA orientation. For example, response time variables can be converted so that shorter times correspond to larger values on a suitable scale, thereby aligning them with the interpretation of outputs in an input-oriented model.

3.4. Interpretation of Efficiency Scores and Decomposition

The DEA results yield both an overall efficiency score for each metro station and separate efficiency scores for the preparedness, response, and recovery stages. These scores are bounded between zero and one under the basic model specification. A score equal to one indicates that the station lies on the efficient frontier and no proportional input reduction is possible without compromising outputs. A score below one implies that inputs could be reduced proportionally while maintaining current performance, or that outputs could potentially be increased given current inputs.

Stage-specific efficiency scores provide insight into which part of the emergency management process contributes most to overall inefficiency. For example, a metro station with relatively high preparedness and response efficiency, but low recovery efficiency, may exhibit sluggish post-earthquake restoration even though its initial management actions are adequate. Conversely, a station with a low preparedness score but higher response and recovery scores may rely heavily on ad hoc efforts and compensatory measures, suggesting that its baseline capacities are insufficient.

Slack variables associated with each input and output further refine the diagnosis. Positive input slacks indicate redundant resource allocation that does not translate into improved performance, while positive output slacks reveal unmet output potential. By examining the pattern of slacks across stages, decision-makers can identify specific

indicators—such as excessive staff deployment, underutilized emergency supplies, or inadequate evacuation coordination—that should be prioritized in improvement strategies.

Efficiency scores can also be classified into qualitative categories to facilitate interpretation, similar to safety level classifications used in other evaluation methods. Metro stations may be labeled as highly efficient, moderately efficient, or inefficient based on predefined threshold ranges. This classification provides a basis for benchmarking and comparative analysis across stations, highlighting best practices as well as critical weaknesses in emergency management.

3.5. Rationale for Method Selection

The choice of a multi-stage DEA model is motivated by both theoretical considerations and practical requirements. Resilience theory emphasizes that system performance under disasters is the result of a sequence of interdependent phases, from preparedness to response and recovery. A multi-stage DEA model naturally aligns with this perspective by explicitly modeling the chained structure of emergency management, rather than aggregating all indicators into a single-stage evaluation.

Compared with traditional multi-criteria methods such as AHP, TOPSIS, and grey relational analysis, DEA does not require subjective weight assignment to inputs and outputs. Instead, it derives an empirical production frontier from observed data, allowing each DMU to be evaluated against the best-performing peers. This characteristic reduces dependence on subjective judgments and enhances the objectivity of the evaluation. At the same time, the network extension of DEA allows for internal structure analysis, so that inefficiencies can be traced to specific stages rather than being treated as an undifferentiated overall deficit.

From a practical perspective, the data requirements of DEA are compatible with the information typically available from metro operation records, emergency planning documents, and expert assessments. The method can accommodate mixed indicator types and does not impose strict distributional assumptions. These features make the multi-stage DEA approach well-suited for evaluating metro station emergency management capability under earthquake scenarios, providing both a rigorous analytical foundation and actionable diagnostic insights.

4. Case Study and Empirical Analysis

To demonstrate the applicability, robustness, and explanatory power of the resilience-oriented multi-stage DEA model established in Section 2 andSection 3, this section presents a detailed empirical study based on Metro Station A. In contrast to simplified examples that rely on composite indicators, the present analysis adheres strictly to the full indicator system defined in Section 2, ensuring that each dimension of absorptive capacity (A1), maintaining capacity (A2), recovery capacity (A3), and adaptive capacity (A4) is explicitly incorporated. This design enhances the integrity of the assessment and strengthens the internal consistency of the methodological framework.

Since existing real-world data for all indicators were not available, scenario-consistent illustrative values were constructed based on expert-informed ranges and commonly reported operational benchmarks in metro emergency management literature. These assumptions do not alter the methodological essence but enable a complete illustration of the evaluation process, from indicator assignment and data transformation through efficiency computation, slack diagnosis, and improvement simulation.

In this study, earthquake scenarios are parameterized in a management-oriented manner rather than through detailed structural damage modeling. Seismic intensity is represented by scenario-based assumptions consistent with local seismic design levels

and emergency planning documents. Damage states are reflected indirectly through their impacts on station functionality, such as equipment availability, communication reliability, evacuation conditions, and recovery duration. These effects are incorporated by adjusting time- and capacity-related indicators, which are subsequently normalized and evaluated within the multi-stage DEA framework.

Due to the limited availability of complete and consistent multi-stage emergency management data across multiple metro stations, this study selects a single representative station (Metro Station A) as an illustrative case. The purpose of this case study is not to produce statistically generalizable efficiency rankings, but to demonstrate the practical applicability of the proposed resilience-oriented multi-stage DEA framework and its ability to diagnose internal process inefficiencies across preparedness, response, and recovery stages. Accordingly, the efficiency results presented in this section should be interpreted as internal performance diagnostics relative to the modeled reference frontier, rather than as comparative measures across different stations.

4.1. Indicator Values and Stage Assignment

The evaluation follows the workflow presented in Section 3. All indicators were first categorized according to the temporal logic of emergency management:

- Preparedness Stage→A1 Indicators. These reflect the station's baseline capacity to resist, absorb, and mitigate the adverse impacts of seismic events.
- Response Stage→A2 Indicators + A1 Output. This stage captures the real-time operational effectiveness immediately after the earthquake occurs.
- Recovery and Adaptation Stage→A3 + A4 Indicators + A2 Outputs. These indicators measure the station's capacity to restore functionality and implement adaptive improvements.

In addition, time-based metrics were normalized using the transformation functions specified in Section 2.2, converting raw measurements into dimensionless efficiency scores compatible with DEA's orientation requirements.

4.1.1. Preparedness Stage Indicators

Preparedness represents the station's fundamental ability to reduce exposure, maintain structural safety, and ensure that emergency resources are adequate and personnel are trained. Table 5 presents the A1 indicators and their assumed values.

Table 5. Preparedness Stage Inputs and Outputs.

Indicator	Symbol	Value	Type
Structural robustness score	$A1_1$	0.82	Input
Equipment anchorage integrity (%)	$A1_2$	0.76	Input
Emergency resource adequacy (%)	$A1_3$	0.70	Input
Staff training completeness (%)	$A1_4$	0.65	Input
Earthquake plan maturity	$A1_5$	0.68	Output

These values collectively illustrate a station that maintains a relatively solid physical foundation but shows deficiencies in personnel readiness and resource deployment; these factors may compromise performance during subsequent stages. The maturity of the emergency plan ($A1_5$) is treated as the output of this stage and acts as an intermediate input for real-time emergency response.

4.1.2. Response Stage Indicators

The response stage evaluates the station's ability to act promptly and effectively once an earthquake occurs. This stage is often the most challenging, as it requires coordination

under time pressure, communication reliability, and adequate staffing. Table 6 shows the indicators used.

Table 6. Response Stage Inputs and Outputs.

Indicator	Symbol	Value	Type
Earthquake plan maturity (from A1)	$A1_5$	0.68	Input
Command response time	$A2_1$	0.40	Output
Communication availability (%)	$A2_2$	0.62	Output
Evacuation coordination effectiveness	$A2_3$	0.55	Output
Subsystem operational continuity (%)	$A2_4$	0.67	Output
Emergency staffing adequacy (%)	$A2_5$	0.58	Output

The assumed values demonstrate that although certain subsystems maintain adequate operational continuity, deficiencies in command response and staffing may substantially impair the speed and effectiveness of emergency actions. These shortcomings often propagate into the recovery phase, influencing restoration times and post-event coordination.

4.1.3. Recovery and Adaptive Stage Indicators

The recovery and adaptation stage evaluates the station's ability to restore critical functions after the earthquake and to enhance long-term resilience through learning and improvement mechanisms. In this stage, recovery capacity indicators ($A3_1$–$A3_5$) are used to measure the efficiency of post-event restoration processes, while adaptive capacity indicators ($A4_1$–$A4_5$) capture institutional learning, plan updating, technological integration, and organizational improvement following the event. All the indicators mentioned above are listed in Table 7.

Table 7. Recovery and Adaptation Stage Inputs and Outputs.

Indicator	Symbol	Value	Type
Response-stage outputs ($A2_1$–$A2_5$)	—	—	Input
Time to restore critical equipment	$A3_1$	0.48	Output
Service recovery level	$A3_2$	0.65	Output
Environmental clearance efficiency	$A3_3$	0.50	Output
Resource deployment timeliness	$A3_4$	0.52	Output
Passenger service restoration time	$A3_5$	0.60	Output
Emergency plan revision frequency	$A4_1$	0.62	Output
Effectiveness of post-event review	$A4_2$	0.60	Output
Improvement in resource allocation mechanisms	$A4_3$	0.58	Output
Integration of new technologies	$A4_4$	0.55	Output
Organizational learning effectiveness	$A4_5$	0.63	Output

These values reflect a station that can restore most service functions within a reasonable timeframe but demonstrates modest performance in adaptive mechanisms that influence long-term resilience enhancement.

4.2. Multi-Stage DEA Evaluation Results

By applying the indicator matrices to the multi-stage DEA model described in Section 3, efficiency values for each stage and for the overall emergency management process were obtained. The results offer a holistic view of the station's performance while revealing critical weaknesses in stage-specific capacities.

4.2.1. Preparedness Stage Efficiency Analysis

Under the proposed multi-stage DEA framework, the preparedness stage exhibits an efficiency score of 0.84 relative to the modeled reference-efficient configuration, indicating that approximately 16% of improvement potential remains within the current preparedness-related inputs and processes. This indicates that while the station maintains a reasonably strong structural and procedural baseline, there remains a performance shortfall of approximately 16% relative to the benchmark frontier.

A detailed examination of slack variables highlights several contributing factors:

(1) Staff training completeness ($A1_4$) exhibits the largest slack, suggesting that existing training programs are insufficient in depth, frequency, or coverage.
(2) Emergency resource adequacy ($A1_3$) also demonstrates notable redundancy, implying either inadequate supplies or suboptimal allocation practices.
(3) Equipment anchorage integrity ($A1_2$), while relatively strong, still falls short of optimal levels under a resilience-oriented evaluation framework.

These results confirm that organizational and resource dimensions, rather than structural conditions alone, significantly constrain preparedness performance.

4.2.2. Response Stage Efficiency Analysis

Under the proposed multi-stage DEA framework, the response stage exhibits an efficiency score of 0.63 relative to the modeled reference-efficient configuration, indicating substantial untapped improvement potential within real-time emergency response processes. This relatively low value reveals substantial operational challenges that directly affect the station's ability to respond to seismic emergencies in real time. Detailed diagnostic findings include:

(1) Command Response Efficiency ($A2_1$). The command response system shows marked inefficiency, suggesting delayed decision-making or communication bottlenecks. Such delays can significantly amplify hazard impacts during the first minutes following an earthquake.
(2) Communication Availability ($A2_2$). Although moderately stable, communication systems do not reach the reliability levels required for high-pressure emergency coordination. Partial subsystem failures or capacity overloads may contribute to this deficiency.
(3) Emergency Staffing Adequacy ($A2_5$). Insufficient on-duty staff undermines evacuation guidance, operational adjustments, and emergency command execution. This factor strongly interacts with $A2_1$ and $A2_3$, creating compounded inefficiencies.
(4) Evacuation Coordination and Subsystem Continuity. While both indicators perform better than $A2_1$ and $A2_5$, their suboptimal scores suggest the need for more refined coordination protocols and better-maintained infrastructure subsystems.

Overall, the response stage constitutes the most critical bottleneck in the entire emergency management chain.

4.2.3. Recovery and Adaptation Stage Efficiency Analysis

Within the constructed resilience-oriented multi-stage DEA framework, the recovery and adaptation stage demonstrates an efficiency score of 0.59 relative to the modeled reference-efficient configuration, suggesting pronounced performance gaps in post-earthquake recovery and adaptation processes.

From the recovery perspective, delays in restoring critical equipment ($A3_1$) and passenger services ($A3_5$) indicate limitations in maintenance preparedness and restoration

coordination. Environmental clearance efficiency ($A3_3$) and resource deployment timeliness ($A3_4$) further constrain the speed of functional recovery.

From the adaptive perspective, moderate scores in emergency plan revision frequency ($A4_1$), effectiveness of post-event review ($A4_2$), and organizational learning effectiveness ($A4_5$) suggest that lessons learned from emergencies are not fully institutionalized. Limited integration of new technologies ($A4_4$) and modest improvements in resource allocation mechanisms ($A4_3$) further restrict long-term resilience enhancement.

In resilience theory, such adaptive components are essential for transforming lessons learned into long-term capability enhancement; their mediocre performance thus restricts the station's ability to evolve toward higher resilience.

4.2.4. Overall Efficiency Assessment

Using the composite evaluation approach detailed in Section 3, the overall efficiency of the emergency management system is 0.55. Under the single-station setting adopted in this study, this value does not represent relative performance against peer stations, but rather reflects internal process efficiency within the proposed multi-stage DEA framework. Specifically, the result indicates that, given the observed intermediate and final outputs, the station could theoretically achieve the same emergency management outcomes with approximately 55% of its current controllable inputs, assuming proportional input reduction across stages and consistency with the internal production structure.

Accordingly, the implied 45% improvement potential should not be interpreted as an immediately realizable or absolute performance gap. Instead, it serves as a diagnostic and directional indicator, highlighting the magnitude of inefficiency embedded in current resource utilization and process coordination. The result underscores the compounding effect of inefficiencies across multiple stages: although preparedness is relatively strong, its benefits are partially offset by shortcomings in the response and recovery processes. This finding reinforces the insight that system-level resilience is constrained by the weakest stage, thereby validating the necessity and analytical value of adopting a multi-stage evaluation framework for metro station emergency management.

4.3. Integrated Diagnosis of Inefficiencies

The multi-stage DEA results enable a comprehensive diagnosis that integrates indicator-level, stage-level, and system-level perspectives.

4.3.1. Preparedness Weaknesses

The largest inefficiencies appear in the organizational and procedural dimensions: (i) Training gaps reduce staff familiarity with emergency protocols, potentially leading to delayed or incorrect actions during an earthquake. (ii) Resource shortages or poor resource management may hinder swift deployment. These weaknesses are likely to propagate into response-stage inefficiencies.

4.3.2. Response-Stage Bottlenecks

The response stage forms the most critical bottleneck, with deficiencies in: Command efficiency, Communication system stability, and Staff availability. Together, these factors impair coordinated action, slow down evacuation, and risk service disruptions. The strong interdependence among response indicators magnifies their negative effects.

4.3.3. Weak Recovery and Limited Adaptive Learning

Recovery efficiency is constrained by Slow restoration of key equipment, Incomplete service recovery, and Moderate clearance efficiency. Meanwhile, adaptive capacity shows structural weaknesses: organizational learning is not adequately institutionalized, and

integration of new technologies remains limited. This indicates that while the station can eventually return to normal operations, its capacity to transform lessons into resilience-enhancing reforms is still insufficient.

4.4. Improvement Simulation and Recalculated Performance

The improvement analysis presented below is designed as a scenario-based what-if simulation aimed at exploring the potential effects of targeted managerial enhancements. The adjusted indicator values do not represent forecasts or guaranteed outcomes, but rather plausible improvement scenarios reflecting short- to medium-term organizational and procedural optimization measures. Based on slack values, targeted improvement measures were applied. The adjustments were designed to reflect realistic, achievable improvements. The improvement of indicators is listed in Table 8, and the Recalculated Efficiency Scores are listed in Table 9.

Table 8. The Indicator improvement.

Indicator	Adjustment
$A1_4$ Staff training completeness	+15%
$A1_3$ Emergency resource adequacy	+10%
$A2_1$ Command response efficiency	+30%
$A2_2$ Communication availability	+20%
$A3_1$ Equipment restoration efficiency	+25%
$A4_1$–$A4_5$ Adaptive capacity indicators	+10% (across-the-board)

Table 9. The Recalculated Efficiency Scores.

Stage	Original	Improved
Preparedness	0.84	0.95
Response	0.63	0.82
Recovery & Adaptation	0.59	0.78
Overall	0.55	0.73

The improvement increments adopted in Table 8 are designed as scenario-based and illustrative adjustments, rather than predictive forecasts or guaranteed outcomes. Their magnitudes are informed by the relative severity of inefficiencies revealed by the DEA slack analysis and by commonly reported ranges of short- to medium-term organizational and procedural improvements in metro emergency management practice. Indicators exhibiting larger inefficiencies and stronger system-wide influence, such as command response efficiency, are therefore assigned higher adjustment levels (e.g., +30%), while indicators associated with more incremental optimization are assigned moderate increases (e.g., +10–20%). Explicit cost estimation is beyond the scope of this methodological study, and the improvement simulation should be interpreted as a what-if analysis exploring plausible enhancement scenarios.

To facilitate practical interpretation, the improvement scenarios in Table 8 should be understood as illustrative translations of efficiency gains into broad categories of emergency management actions, rather than prescriptive implementation plans. For example, a "+30% improvement in command response efficiency" reflects enhanced effectiveness in decision-making and coordination, which may be associated with streamlined command procedures, clearer decision protocols, or improved decision-support tools that reduce response latency in the early post-earthquake period. Similarly, a "+20% improvement in communication availability" represents improved robustness of communication functions under emergency conditions, potentially linked to redundancy planning, system testing, or

enhanced staff familiarity with degraded-mode communication protocols. Improvements in equipment restoration efficiency and resource deployment timeliness can be interpreted as indicative of more effective recovery organization and coordination, rather than specific engineering interventions.

These examples are intended to support the conceptual interpretation of the improvement simulation results, illustrating how the DEA-based diagnostic outcomes relate to plausible organizational, procedural, and technological enhancement directions at the station level, without implying guaranteed performance gains or detailed cost assumptions.

The improvement simulation highlights that strategic enhancements—especially in training, communication, and recovery procedures—can significantly uplift the station's resilience performance. Most notably, improvements in preparedness produce cascading benefits across subsequent stages, validating the system-thinking perspective embedded in the resilience-oriented DEA framework.

4.5. Summary

Through a rigorous and systematic multi-stage DEA assessment, this section provides a comprehensive evaluation of Metro Station A's emergency management performance under earthquake scenarios. The findings confirm that:

- Preparedness is solid but not free of weaknesses, notably in training and resource readiness.
- Response represents the most critical bottleneck, with command efficiency and communication reliability emerging as dominant limiting factors.
- Recovery and adaptive processes require deeper institutionalization, particularly in equipment restoration and organizational learning.
- The overall efficiency remains modest, reflecting the cumulative effect of inefficiencies across all stages.
- Targeted improvement strategies can yield substantial gains, raising overall efficiency from 0.55 to 0.73.

The section demonstrates the analytical power of the resilience-oriented multi-stage DEA model and provides a clear, actionable framework for improving metro emergency management performance. While the empirical application in this study focuses on a single metro station to illustrate the methodological validity and diagnostic capability of the proposed framework, the model is inherently scalable. With appropriate data availability, the framework can be readily extended to a multi-station dataset, enabling comparative efficiency analysis across stations and further exploration of scale and structural heterogeneity.

However, when extending the framework to multiple stations, it should be noted that DEA-based efficiency analysis is subject to minimum sample size requirements to avoid overfitting and artificial efficiency inflation. In practical applications, the number of decision-making units (DMUs) should be sufficiently large relative to the total number of inputs and outputs, following commonly adopted rules of thumb in the DEA literature (e.g., $n \geq \max\{m \times s, 3(m+s)\}$). When the sample size is limited, indicator aggregation or dimensionality-reduction strategies may be required to ensure robust efficiency estimation.

5. Conclusions

This study proposed a resilience-oriented framework for evaluating metro station emergency management under earthquake conditions. By integrating a multi-dimensional resilience indicator system with a multi-stage DEA model, the research conceptualized emergency management as a sequential managerial process involving preparedness, response, and recovery. The empirical analysis demonstrated that this approach is capable of revealing not only overall performance but also stage-specific managerial weaknesses that traditional evaluation methods often overlook. From a management perspective, the results

show that many performance gaps stem not from infrastructural limitations but from managerial processes such as training effectiveness, communication reliability, decision-making efficiency, and coordination during restoration. The improvement simulation further confirms that strengthening these managerial components can significantly enhance overall emergency efficiency, underscoring the central role of management capability in achieving seismic resilience.

The study makes several theoretical contributions to the management field. It operationalizes resilience as a measurable, multi-stage managerial construct rather than a static technical attribute, and demonstrates how efficiency analysis can be applied to complex organizational processes. The integration of resilience theory with network DEA extends existing emergency management research by providing a process-oriented, data-driven method for diagnosing inefficiencies within and across organizational stages. Practical recommendations arising from the findings include prioritizing scenario-based staff training, establishing redundant and robust communication systems, formalizing post-event recovery protocols, and strengthening organizational learning mechanisms. These measures directly target the managerial factors shown to have the greatest influence on emergency performance. Although the empirical application in this study focuses on a single metro station, this design enables an in-depth, process-level diagnosis of emergency management performance that would be difficult to achieve using highly aggregated or purely cross-sectional datasets. The proposed framework is inherently scalable and can be readily extended to multiple stations or network-level analyses when more comprehensive data becomes available. Future research should expand the dataset, explore inter-organizational coordination mechanisms, and incorporate advanced modeling techniques to capture dynamic managerial behaviors.

In summary, this research reinforces that resilience in metro systems is fundamentally a management challenge. By providing a structured approach for assessing, comparing, and improving emergency management performance, the proposed framework supports evidence-based decision-making and contributes to the broader understanding of how organizations can better prepare for, respond to, and recover from disruptive events.

Author Contributions: Conceptualization, L.Z. and H.Y.; methodology, L.Z. and H.Y.; formal analysis, L.Z.; investigation, L.Z.; data curation, L.Z.; writing—original draft preparation, L.Z.; writing—review and editing, H.Y.; visualization, L.Z.; supervision, H.Y.; project administration, H.Y.; funding acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the Science and Technology Plan Project of Longquanyi District of Chengdu City (No. 2025LQRD0030), administered by the New Economy and Science and Technology Bureau of Longquanyi District, Chengdu City. The funds were used for sample collection, analysis, and statistical software licenses. The funding agencies had no influence on the study design, data collection, interpretation of results, or the writing of this manuscript.

Data Availability Statement: The datasets analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chen, E.; Liu, Y.; Yang, M.; Ye, Z.; Nie, Y.M. The Sustainability Appeal of Urban Rail Transit. *Transp. Res. Part A Policy Pract.* **2024**, *186*, 104152. [CrossRef]
2. Hu, J.; Wen, W.; Zhai, C.; Pei, S. A Comprehensive Review of Resilience of Urban Metro Systems: A Perspective from Earthquake Engineering. *Tunn. Undergr. Space Technol.* **2024**, *152*, 105920. [CrossRef]

3. Wei, Y.; Yang, X.; Xiao, X.; Ma, Z.; Zhu, T.; Dou, F.; Wu, J.; Chen, A.; Gao, Z. Understanding the Resilience of Urban Rail Transit: Concepts, Reviews, and Trends. *Engineering* **2024**, *41*, 7–18. [CrossRef]
4. Bommer, J.J. Earthquake Hazard and Risk Analysis for Natural and Induced Seismicity: Towards Objective Assessments in the Face of Uncertainty. *Bull. Earthq. Eng.* **2022**, *20*, 2825–3069. [CrossRef]
5. Yoo, M.; Park, J.; Jeong, J.; Ahn, S. Analysis of Influence of Seismic Impact on Evacuability in Subway Stations. *Dev. Built Environ.* **2025**, *21*, 100637. [CrossRef]
6. Zito, M.; Nascimbene, R.; Dubini, P.; D'Angela, D.; Magliulo, G. Experimental Seismic Assessment of Nonstructural Elements: Testing Protocols and Novel Perspectives. *Buildings* **2022**, *12*, 1871. [CrossRef]
7. Dong, Z.; Li, C.; Wen, S.; Zhai, P.; Kang, S. Study on Seismic Resilience Quantitative Framework of Subway Stations and Resilience Improvement Strategy. *Front. Earth Sci.* **2022**, *10*, 869965. [CrossRef]
8. Zhang, P.; Peeta, S. A Generalized Modeling Framework to Analyze Interdependencies among Infrastructure Systems. *Transp. Res. Part B Methodol.* **2011**, *45*, 553–579. [CrossRef]
9. Ouyang, M.; Wang, Z. Resilience Assessment of Interdependent Infrastructure Systems: With a Focus on Joint Restoration Modeling and Analysis. *Reliab. Eng. Syst. Saf.* **2015**, *141*, 74–82. [CrossRef]
10. Song, J.; Miao, H.; Zhong, Z.; Zhao, M.; Du, X. A Novel Probabilistic Method for Assessing the Seismic Resilience of Subway Systems across Four Dimensions. *Int. J. Disaster Risk Reduct.* **2025**, *131*, 105898. [CrossRef]
11. Bruneau, M.; Chang, S.E.; Eguchi, R.T.; Lee, G.C.; O'Rourke, T.D.; Reinhorn, A.M.; Shinozuka, M.; Tierney, K.; Wallace, W.A.; von Winterfeldt, D. A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. *Earthq. Spectra* **2003**, *19*, 733–752. [CrossRef]
12. Cimellaro, G.P.; Reinhorn, A.M.; Bruneau, M. Framework for Analytical Quantification of Disaster Resilience. *Eng. Struct.* **2010**, *32*, 3639–3649. [CrossRef]
13. Kammouh, O.; Chahrour, N. Indicator-Based Framework to Evaluate the Resilience of Transport Infrastructure Systems. *Sustain. Resilient Infrastruct.* **2025**, *10*, 536–562. [CrossRef]
14. Tian, T.; Liang, Y.; Peng, Z.; Cheng, Y.; Chen, K. Assessing the Dynamic Resilience of Urban Rail Transit Networks during Their Evolution Using a Ridership-Weighted Network. *PLoS ONE* **2023**, *18*, e0291639. [CrossRef] [PubMed]
15. Jiang, J.; El Naggar, M.H.; Huang, W.; Xu, C.; Zhao, K.; Du, X. Seismic Vulnerability Analysis for Shallow-Buried Underground Frame Structure Considering 18 Existing Subway Stations. *Soil Dyn. Earthq. Eng.* **2022**, *162*, 107479. [CrossRef]
16. Tsinidis, G.; Karatzetzou, A.; Stefanidou, S.; Markogiannaki, O. Developments in Seismic Vulnerability Assessment of Tunnels and Underground Structures. *Geotechnics* **2022**, *2*, 209–249. [CrossRef]
17. Cui, Z.-D.; Zhang, L.-J.; Zhan, Z.-X. Seismic Response Analysis of Shallowly Buried Subway Station in Inhomogeneous Clay Site. *Soil Dyn. Earthq. Eng.* **2023**, *171*, 107986. [CrossRef]
18. Liu, Y.; Zhuang, H.; Zhang, J.; Liu, Z. Seismic Performance of Underground Subway Station Structure near the Strike-Slip Fault by the Developed Hybrid IBE-FEM Method. *Eng. Anal. Bound. Elem.* **2024**, *166*, 105851. [CrossRef]
19. Wen, W.; Hu, J.; Zhai, C.; Pei, S. Post-Earthquake Functionality Assessment of Subway Stations Considering the Interdependency among Sub-Systems. *Comput.-Aided Civ. Infrastruct. Eng.* **2024**, *39*, 136–161. [CrossRef]
20. Xu, P.-C.; Lu, Q.-C.; Feng, T.; Li, J.; Li, G.; Xu, X. Resilience Analysis of Metro Stations Integrating Infrastructures and Passengers. *Reliab. Eng. Syst. Saf.* **2024**, *252*, 110467. [CrossRef]
21. Zhang, D.; Zhang, C.; Lei, C.; Huang, Z.; Huang, H. Multi-Parameter Assessment Method for Seismic Resilience of Subway Stations Based on Economic Losses and Recovery Paths. *Front. Struct. Civ. Eng.* **2025**, *19*, 782–807. [CrossRef]
22. Zhou, Y.; Chen, J.; Zhong, M.; Hua, F.; Sui, J. Evacuation Effect Analysis of Guidance Strategies on Subway Station Based on Modified Cellular Automata Model. *Saf. Sci.* **2023**, *168*, 106309. [CrossRef]
23. Shan, S.; Guo, X.; Wei, Z.; Sun, W.; Zheng, H.; Pan, H.; Lin, J. Simulation Analysis of Evacuation Processes in a Subway Station Based on Multi-Disaster Coupling Scenarios. *Int. J. Disaster Risk Reduct.* **2023**, *96*, 103998. [CrossRef]
24. Tang, Y.; Xia, N.; Lu, Y.; Varga, L.; Li, Q.; Chen, G.; Luo, J. BIM-Based Safety Design for Emergency Evacuation of Metro Stations. *Autom. Constr.* **2021**, *123*, 103511. [CrossRef]
25. Tang, Y.; Bi, W.; Varga, L.; Dolan, T.; Li, Q. An Integrated Framework for Managing Fire Resilience of Metro Station System: Identification, Assessment, and Optimization. *Int. J. Disaster Risk Reduct.* **2022**, *77*, 103037. [CrossRef]
26. Ding, Z.; Xu, S.; Xie, X.; Zheng, K.; Wang, D.; Fan, J.; Li, H.; Liao, L. A Building Information Modeling-Based Fire Emergency Evacuation Simulation System for Large Infrastructures. *Reliab. Eng. Syst. Saf.* **2024**, *244*, 109917. [CrossRef]
27. Hui, Y.; Su, S.; Peng, H. Evaluation of Subway Emergency Evacuation Based on Combined Theoretical and Simulation Methods. *Appl. Sci.* **2024**, *14*, 11580. [CrossRef]
28. Liu, J.; Qi, Y.; Wang, W. Emergency Management Capacity Assessment for Urban Rail Transit—An Example of Beijing Metro Line 13. *Transp. Saf. Environ.* **2024**, *6*, tdad015. [CrossRef]
29. Chu, J.; Zhu, J. Production Scale-Based Two-Stage Network Data Envelopment Analysis. *Eur. J. Oper. Res.* **2021**, *294*, 283–294. [CrossRef]

30. Lim, D.-J.; Kim, M.-S. Measuring Dynamic Efficiency with Variable Time Lag Effects. *Omega* **2022**, *108*, 102578. [CrossRef]
31. Ratner, S.V.; Shaposhnikov, A.M.; Lychev, A.V. Network DEA and Its Applications (2017–2022): A Systematic Literature Review. *Mathematics* **2023**, *11*, 2141. [CrossRef]
32. Rathnayaka, B.; Robert, D.; Adikariwattage, V.; Siriwardana, C.; Meegahapola, L.; Setunge, S.; Amaratunga, D. A Unified Framework for Evaluating the Resilience of Critical Infrastructure: Delphi Survey Approach. *Int. J. Disaster Risk Reduct.* **2024**, *110*, 104598. [CrossRef]
33. Bruckler, M.; Wietschel, L.; Messmann, L.; Thorenz, A.; Tuma, A. Review of Metrics to Assess Resilience Capacities and Actions for Supply Chain Resilience. *Comput. Ind. Eng.* **2024**, *192*, 110176. [CrossRef]
34. Ge, S.; Shan, M.; Zhai, Z. Emergency Preparedness in China's Urban Rail Transit System: A Systematic Review. *Sustainability* **2025**, *17*, 524. [CrossRef]
35. Zhang, J.; Huang, D.; You, Q.; Kang, J.; Shi, M.; Lang, X. Evaluation of Emergency Evacuation Capacity of Urban Metro Stations Based on Combined Weights and TOPSIS-GRA Method in Intuitive Fuzzy Environment. *Int. J. Disaster Risk Reduct.* **2023**, *95*, 103864. [CrossRef]
36. Dui, H.; Guo, W.; Xia, W.; Wu, S. Digital Twin-Based Resilience Analysis and Emergency Maintenance with Generative AI of Smart Urban Metro Systems. *Reliab. Eng. Syst. Saf.* **2026**, *265*, 111537. [CrossRef]
37. Xing, J.; Yin, X.; Zhang, J.; Chen, J. Resilience Modeling and Improvement of Metro Systems Considering Statistical Behaviors of Passenger Mobility. *Int. J. Disaster Risk Reduct.* **2023**, *96*, 103975. [CrossRef]
38. Li, C.; Wang, Y.; Qing, W.; Li, C. What Makes a Specialised Emergency Plan Effective?—A Case Study of Railway Operation in 33 Prefectural Cities in China. *PLoS ONE* **2025**, *20*, e0327007. [CrossRef]
39. Fukuyama, H.; Tsionas, M.; Tan, Y. Dynamic Network Data Envelopment Analysis with a Sequential Structure and Behavioural-Causal Analysis: Application to the Chinese Banking Industry. *Eur. J. Oper. Res.* **2023**, *307*, 1360–1373. [CrossRef]
40. Chen, L.; Wang, Y.-M. Efficiency Decomposition and Frontier Projection of Two-Stage Network DEA under Variable Returns to Scale. *Eur. J. Oper. Res.* **2025**, *322*, 157–170. [CrossRef]
41. Afonso, G.P.; Ferreira, D.C.; Figueira, J.R. A Network-DEA Model to Evaluate the Impact of Quality and Access on Hospital Performance. *Ann. Oper. Res.* **2024**, *342*, 2169–2199. [CrossRef]
42. Long, Q.; Song, K. Operational Performance Evaluation of E-Government Microblogs Under Emergencies Based on a DEA Method. *Inf. Syst. Front.* **2022**, *24*, 1–18. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.