Jurnal Presipitasi
Media Komunikasi dan Pengembangan Teknik Lingkungan
e-ISSN: 2550-0023

Vol 23, No 1, 2025,
Article available at homepage presipitasi

Regional Case Study

Decision-Ready Composite Performance Index for Raw Water Supply Systems: PLS-SEM and GRG

Ussy Andawayanti[1]*, Ery Suhartanto[1], Rahmah Dara Lufira[1], Hari Siswoyo[1], Sri Utami Sudiarti[1], Rizki Ramadhani Pratama[2]

[1] Department of Water Resources Engineering, Universitas Brawijaya, Jalan MT. Haryono, Malang, Indonesia 65145

[2] Department of Civil and Environmental Engineering, Nagoya University, Japan

*Corresponding Author, email: uandawayanti@ub.ac.id

Abstract

Dependable raw-water service depends on asset condition, institutional capability, and watershed context, existing checklists in Indonesia fail to produce a validated, decision-ready performance score. This study develops a composite performance indicator for raw water infrastructure that incorporates technical (Tk), institutional/non-technical (NT), and environmental (Li) dimensions. Data were collected from 21 schemes in Lombok–Sumbawa, West Nusa Tenggara Province, Indonesia (NTB), with 160 respondents, using field assessments and 1–4 scale questionnaires. Estimated reflective formative PLS-SEM, then applied GRG calibration to minimize deviation from field scores under non-negativity and unit sum constraints for interpretability and portability. All pillars contribute positively and significantly to the composite index, which exhibits high explanatory power ($R^2 = 0.997$). The calibrated index is PIRWSS = 0.440 PI_{Tk} + 0.340 PI_{NT} + 0.220 PI_{Li}, with SSR ≈ 83.412, RMSE ≈ 0.522, MSE ≈ 5.721, and ≈ 99.70% accuracy relative to field benchmarks. Cross-site analysis shows higher performance in Lombok than in Sumbawa, reflecting hydroclimatic conditions and conveyance configurations. The index provides utilities and regulators with a transparent, reproducible framework for benchmarking and prioritizing operations, maintenance, rehabilitation, and source-water protection.

Keywords: Raw water supply systems; performance index for raw water supply systems (pirwss); pls-sem; grg optimization; watershed management

1. Introduction

Reliable raw water supply underpins public health, economic productivity, and the resilience of downstream drinking water services, yet performance is difficult to appraise because raw water systems are socio-technical networks spanning source catchments, intake works, transmission pipelines, treatment interfaces, and the institutions that operate them. In Indonesia's island provinces, such as West Nusa Tenggara Province, Indonesia (NTB), hydrologic seasonality, land-use change, and dispersed settlements amplify operational risks, heightening the need for an empirically grounded and operationally usable assessment framework to guide maintenance, rehabilitation, and investment (Guo et al., 2020). Current national practice references the Directorate General of Water Resources Circular Letter (SE) 03/2021, a readiness for operation and maintenance checklist adapted from irrigation performance concepts. While useful for documenting infrastructure elements and organizational arrangements, SE 03/2021 omits key environmental attributes of source areas and does not yield a quantified, system-level performance index, consequently, gaps persist between ideal requirements and field realities. Evidence from NTB indicates that several variables needed to reflect actual raw water performance particularly environmental and non-technical dimensions are either insufficiently in assessments now in use (Zakiyayasin Nisa' et al., 2023).

Internationally, two strands dominate the assessment landscape. Utility oriented systems emphasize service continuity, losses, adequacy, and quality, but overwhelmingly target distribution rather than the upstream raw water segment and often rely on expert judgment rather than reproducible, data driven aggregation. The International Water Association (IWA) recommends KPIs coverage, quality compliance, and efficiency that, while valuable, require tailoring to raw water functions (Moudi, 2022). In parallel, urban water security and integrated water resources management frameworks broaden the lens to governance and basin conditions, including Sustainable Development Goal 6, yet typically operate at city or basin scales too coarse to guide day-to-day operation and maintenance (O & M), decisions for specific intakes and conveyance systems (Engelenburg *et al.*, 2021). Consequently, asset rehabilitation and O&M prioritization for raw water systems can be poorly aligned with actual performance constraints.
Environmental change further strengthens the case for integrating watershed indicators into raw water performance assessment. Deforestation, agricultural intensification, and urban expansion can degrade source quality and reliability, increasing turbidity and nutrient loads and driving up treatment costs; climate variability compounds these pressures and introduces greater uncertainty in seasonal availability (Macharia et al., 2021). Performance tools that ignore environmental condition therefore risk misdiagnosing bottlenecks and misprioritizing investments.

Composite indicator design offers a transparent pathway to integrate heterogeneous evidence technical, institutional, and environmental into a single decision aid when built on clear rules for variable selection, normalization, weighting, and validation. Latent variable modeling provides the statistical machinery to bind these dimensions, variance-based structural equation modeling (PLS-SEM) is well suited when constructs are measured by multiple indicators, sample sizes are modest, and distributional assumptions are relaxed. Established criteria ensure measurement quality while a structural model estimates the relative contribution of each latent dimension to overall performance (Joseph F Hair et al., 2017). Numerical optimization can then refine weights to improve predictive fit under practical constraints; the generalized reduced gradient (GRG) method remains a robust choice for constrained nonlinear problems of this kind (Lasdon et al., 1978). Empirically, across 21 raw-water schemes in NTB with inputs from 160 stakeholders, we document quantifiable constraints: an intaketransmission line designed for 50 L/s delivers only 30 L/s at the treatment plant (idle capacity 20 L/s); Lombok's wetter, largely gravity-fed context (1,441 mm/year) contrasts with drier, pump-dependent Sumbawa (1,176 mm/year), and several Sumbawa schemes fall in "poor/less-good" categories, evidence of real performance deficits rather than assumptions. These conditions elevate O&M burdens and reliability risks, making a decision-ready index for the raw-water segment operationally urgent (BWS NT I, 2023). Against this backdrop, a clear gap persists. Most utility metrics prioritize distribution-side outcomes, and most water security frameworks aggregate at scales too coarse for O&M, few instruments directly target the raw water segment as a distinct operational domain linking source waters to treatment plants. Even fewer combine a validated latent-variable structure with numerical optimization and verify predictive performance against field data across diverse hydro-climatic settings (Chakraborty et al., 2022).

This study addresses those gaps by proposing and validating a composite Performance Index for Raw Water Supply Systems (PIRWSS) that integrates three pillars technical, non-technical (institutional/managerial), and environmental so that asset condition, institutional capacity, and catchment characteristics jointly inform an overall score. The research was undertaken on Lombok and Sumbawa (NTB) using field surveys and questionnaires administered to 160 stakeholders across 21 locations, covering a diversity of source types, conveyance configurations, and managerial contexts. Methodologically, we specify a reflective formative measurement structure and estimate a structural model using PLS-SEM, then calibrate coefficients via GRG subject to normalization and monotonicity constraints to enhance predictive fidelity. The novelty lies in extending SE 03/2021 with an explicit, quantified environmental pillar, combining validated latent variable modeling with GRG calibration to produce reproducible weights, and demonstrating predictive accuracy across multiple real world sites in an island context where environmental pressures and institutional capacity jointly shape outcomes.

Objectives are to: (i) evaluate existing systems using SE 03/2021 as baseline evidence; (ii) quantify relationships among technical, non-technical, and environmental dimensions via PLS-SEM; (iii) calibrate PIRWSS weights via GRG under interpretability constraints; (iv) validate predictive alignment against field scores across 21 schemes; and (v) provide a transparent benchmark to prioritize O&M, rehabilitation, program handover, and policy dialogue. The impact is practical and immediate, the PIRWSS enables comparable, site specific benchmarking, supports O&M planning, rehabilitation, and program handover, and provides a transparent basis for policy dialogue on sustainability oriented management of raw water supply systems in NTB and similar settings.

2. Methods

2.1. Study Area and Data Used

The study focuses on the raw-water segment in Lombok and Sumbawa, West Nusa Tenggara Province, Indonesia (NTB), from source catchments and intake works through gravity- or pump-fed conveyance to pre-treatment reservoirs, i.e., upstream of drinking water distribution. Twenty-one sites spanning heterogeneous hydro-climatic and topographic settings were surveyed to capture variation in source types, conveyance configurations, and organizational contexts. A mixed-methods design combined primary and secondary evidence. Primary data comprised site visits, semi-structured observations, structured interviews, documentation, and questionnaire surveys administered to utility operators, provincial planners, river-basin staff, and water users. Two sampling waves yielded 160 respondents (wave-1 = 111; wave-2 = 49). Instruments used a 1–4 scale with indicator-specific descriptors to maintain alignment between field observation and respondent scoring; all instruments were piloted and refined to improve clarity and reduce measurement error. Secondary sources included technical inventories (asset registers, production logs), institutional records (budgets, staffing, SOPs), and hydro-environmental information (land cover, watershed condition). As a sectoral baseline, the national technical guidance for raw-water performance appraisal (SE 03/2021) was reviewed alongside administrative and technical records to cross-check responses and reconcile inconsistencies.

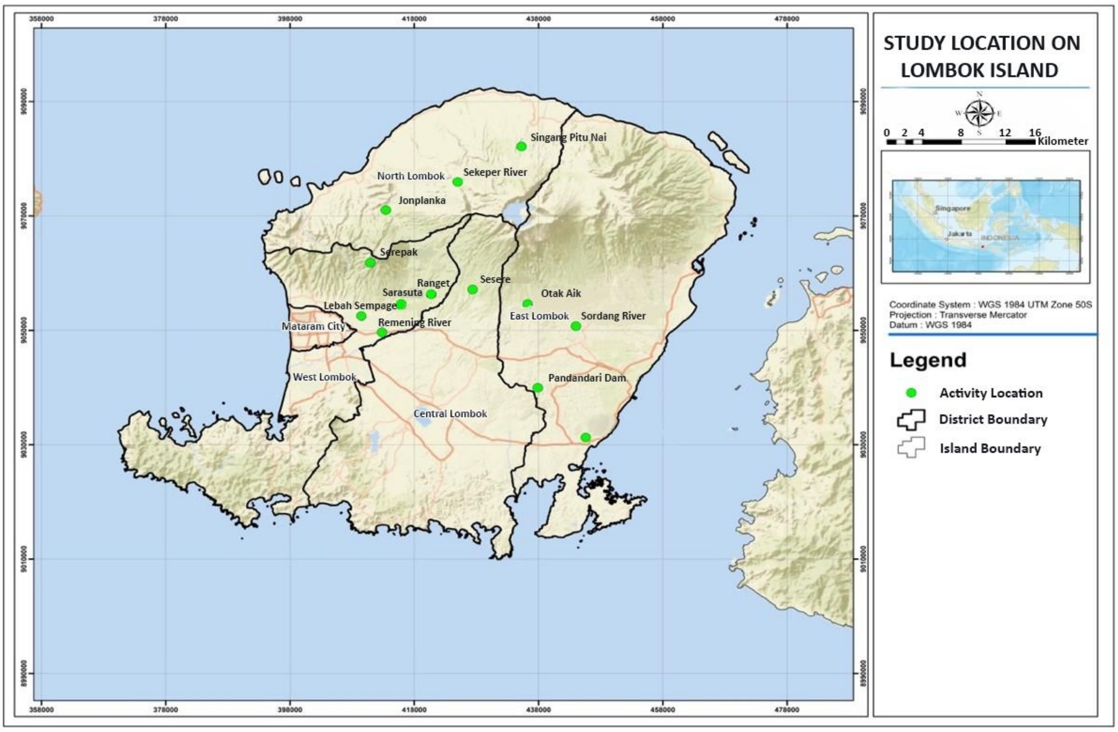

Figure 1. Study locations on Lombok Island

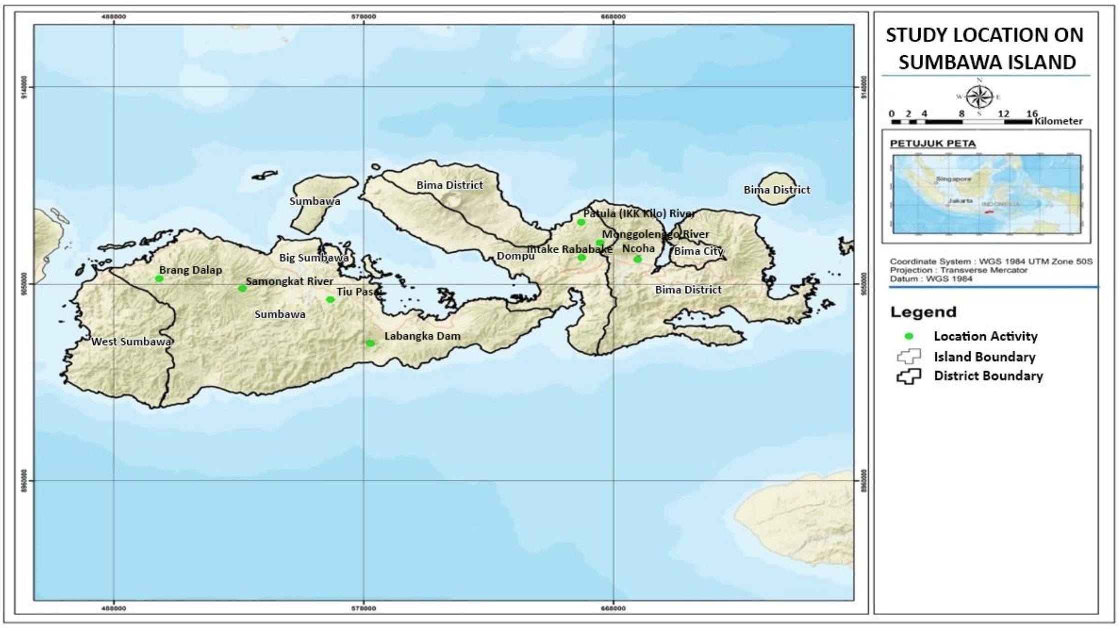

Figure 2. Study locations on Sumbawa Island

Figure 1 and Figure 2 illustrate the research site in connection with Table 1 below, detailing areas throughout NTB, including Lombok and the Sumbawa Islands, comprising 21 research sites. Thirteen spots on Lombok Island and eight locations on Sumbawa Island have all been administered by the General Institution of Drinking Water Authority in Indonesia (PDAM).

Table 1. Research location database raw water

No.	Unit of raw water	Regency	Type of source	Type of intake structure
Lombok Island				
1	Lebah Sempage	West Lombok	Water source	Broncaptering
2	Sarasuta	West Lombok	Water source	Broncaptering
3	Remening	West Lombok	River	Intake of weir
4	Serepak	West Lombok	River	Intake of weir
5	Sesera	Central Lombok	Water source	Broncaptering
6	Rangat	West Lombok	Water source	Broncaptering
7	Pandanduri	East Lombok	Dam	Intake of dam
8	Sekeper	North Lombok	River	Intake of weir
9	Tibu Ulik	East Lombok	Small dam	Intake of small dam
10	Sordang	East Lombok	River	Free Intake
11	Singang Pitu Nai	North Lombok	River	Free Intake
12	Jonplanka	North Lombok	Water source	Broncaptering
13	Otak Aik	West Lombok	Water source	Broncaptering
Sumbawa Island				
1	Semongkat	Sumbawa	River	Intake of weir
2	Brangdalap	Sumbawa	River	Intake of weir
3	Tiu Pasai	Sumbawa	Small dam	Intake of small dam
4	Labangka	Sumbawa	Dam	Intake of dam
5	Monggelenggo	Dompu	River	Intake of weir
6	Ncoha	Bima	Small dam	Intake of small dam

No.	Unit of raw water	Regency	Type of source	Type of intake structure
7	Patula	Bima	River	Intake of weir
8	Rababaka	Dompu	River	Intake of weir

2.2. Baseline Assesment Framework

Directorate General of Water Resources Circular SE 03/2021 structures appraisal into six variables physical infrastructure, service productivity, supporting facilities, institutions and human resources, documentation, and water-user associations with predefined weights. Gap analysis revealed two shortcomings for decision support in NTB, several regulatory variables are not operationally tied to measurable field conditions and environmental/source area indicators are under-represented despite their known influence on reliability and treatment effort. A variable mapping matrix therefore aligned the six regulatory variables with three higher-level latent dimensions used in the model: technical, non-technical, and environmental (Mergoni, Inverno and Carosi, 2022).

The technical dimension was operationalized by two parameter groups: (A1) source quantity and reliability (availability, production capacity, continuity, losses, O&M adequacy) and (A2) physical asset condition (intakes, conveyance mains, supporting works). The non-technical dimension encompassed governance and managerial readiness (budgeting, staffing, planning, documentation, user associations). The environmental dimension incorporated watershed condition (forest cover, open land, plantation/agriculture, built-up area) and source water quality proxies. Indicators were defined with transparent 1–4 descriptors to minimize rater ambiguity and ensure reproducibility (Table 2). To preserve diagnostic value and reduce compensability, aggregation followed a hierarchical scheme, indicators, parameters, dimensions, composite index, with normalization via ratio or ordinal scaling as appropriate. Missing values were handled via listwise deletion when infrequent and via conservative imputation otherwise. For external comparability with regional utility practice, baseline scoring referenced the four aspect drinking water supply system improvement agency in Indonesia (BPPSPAM) structure financial (0.25), service (0.25), operational (0.35), human resources (0.15) as a non-physical benchmark for later validation. Crucially, weights at each aggregation tier were estimated empirically rather than assigned a priori, consistent with composite-indicator best practice (Mao et al., 2019).

Table 2. Indicator hierarchy and measurement model,

Description	Performance Value (%)	Performance Value (Number)
Very good performance	80-100	4
Good performance	70-80	3
Not good performance	55-70	2
Bad performance	<55	1

Source: Directorate General of Water Resources Circular SE 03/2021

2.3. PLS-SEM Model

Given multi-indicator constructs and a modest sample, we employed variance-based structural equation modeling (PLS-SEM) to estimate measurement and structural components. Indicators were specified reflectively within the technical, non-technical, and environmental dimensions. Measurement quality followed established criteria, indicator reliability (outer loadings ideally ≥ 0.70), internal consistency (Composite Reliability ≥ 0.70), convergent validity (AVE ≥ 0.50), and discriminant validity via the Fornell–Larcker criterion and the heterotrait monotrait ratio (Fornell et al., 1981). Low-loading indicators were pruned only when justified by both statistics and field plausibility; the final run satisfied AVE, CR, and α benchmarks. The inner model related the three pillars to the composite performance index PIRWSS, with path coefficients obtained via bootstrapping. Collinearity was examined through inner VIFs; predictive relevance was assessed using Stone–Geisser's Q^2 and, where feasible, cross-validated redundancy (Joseph F Hair et al., 2017). This SEM stage yielded a closed-form estimator of the composite index and provided data-driven pillar weights for subsequent calibration. Best-practice considerations for infrastructure index construction guided the two-step evaluation (measurement to structural) and reporting of reliability, R^2, effect sizes (f^2), and pathway significance (Alismaiel, 2021; Tefera and Hunsaker, 2021).

2.4. Calibration and Validation Generalized Reduced Gradient (GRG)

To enhance predictive alignment with observed field performance, we calibrated weights using the Generalized Reduced Gradient (GRG) method, a robust approach for constrained nonlinear optimization (Lasdon et al., 1974, 1978). The objective minimized the sum of squared residuals (SSR) between modeled and observed site-level performance subject to non-negativity and unit-sum constraints on weights at each tier; monotonicity (higher indicator values must not reduce the composite); and stability penalties to avoid extreme, non-interpretable weights. Calibration was performed at sub-index and composite levels to maintain coherence between measurement and decision layers. In solver settings and reduced-gradient diagnostics were documented; illustrative outputs for the technical sub-index report RMSE and SSR traces used to confirm convergence. The GRG step is complementary to PLS-SEM refining empirically estimated coefficients within practical constraints to produce a decision-ready index (Joe F Hair et al., 2017). Related optimization guidance in composite-index contexts further supports this integration (Budianto et al., 2025; Sankar, 2024).

Model quality was examined on multiple fronts. In-sample statistical validity drew on PLS-SEM diagnostics construct reliability/validity (α, CR, AVE, HTMT), explained variance (R^2), and path significance to confirm adequacy of measurement and structural specifications (Fornell et al., 1981). Predictive validation compared modelled PIRWSS against field-measured performance for all 21 locations; the calibrated model closely reproduced observed values, with consistent classification under the study's performance classes. Content validity was established via expert review in stakeholder meetings, ensuring indicators and weights were credible to practitioners and aligned with operational realities. To guard against overfitting, we conducted sensitivity analyses to perturb indicator values and weights within plausible bounds and observed rank stability across sites. Where data allowed, hold-out tests and cross-validated redundancy (Q^2) were used to assess out-of-sample relevance (Joseph F Hair et al., 2017). Environmental indicators were triangulated with independent, map-based land-cover evidence using a standardized legend, reducing observer bias and reflecting contemporary links between watershed condition, reliability, and treatment effort. For transparency and replication, the dissertation provides worked detailing computation of pillar scores and PIRWSS using the calibrated weights, alongside the classification thresholds used operationally for raw-water networks.

A composite-index design was adopted because it converts heterogeneous technical, institutional, and environmental indicators into a single decision aid that supports benchmarking and policy dialogue an approach widely used in water and nexus assessments and recommended to accompany transparency in normalization, weighting, and sensitivity analysis (Ayadi et al., 2024; He et al., 2024; Simpson et al., 2022). PLS-SEM was chosen to estimate a prediction-oriented latent structure with multi-indicator constructs under modest sample sizes and relaxed distributional assumptions, consistent with our field design and prior guidance (Joseph F Hair et al., 2017). Constrained numerical calibration via GRG was then used to align modeled scores with observed site performance while enforcing interpretability (non-negativity, unit-sum, monotonicity); optimization-aided index construction is increasingly applied when aggregating multi-source, multi-scale evidence in water performance monitoring to improve robustness and decision usefulness (Su and Cao, 2022). This sequencing (PLS-SEM and GRG calibration) thus reflects the data structure (n=21 sites; mixed measurement scales), operational constraints, and the study's goal of producing a transparent, decision-ready index for O&M (Lasdon et al., 1974). See also recent utility benchmarking frameworks showing the value of composite performance layers for managerial action (Ganjidoost et al., 2021).

3. Result and Discussion

3.1. Measurement and Structural Model (PLS-SEM) and Generalized Reduced Gradient (GRG)

The study encompassed 21 raw-water supply sites in NTB, 13 in Lombok and 8 in Sumbawa capturing contrasts in hydro-climatic regimes and infrastructure typologies. Field surveys and questionnaires yielded 160 respondent records from utility staff, basin agencies, and users, providing observed-condition data and structured judgments for indicator scoring, site narratives also record event-driven variability most notably the 2018 earthquake and subsequent rehabilitation reinforcing the need to couple technical and environmental diagnostics within routine performance assessment. Variance-based PLS–SEM supported a three-pillar structure technical (PI_{Tk}), non-technical (PI_{NT}), and environmental (PI_{Li}). Measurement quality met accepted criteria for reflective models: high outer loadings, CR ≥ 0.70, α ≥ 0.70, and AVE ≥ 0.50 (Table 3); discriminant validity was confirmed with cross-loadings highest on their intended constructs (Fornell et al., 1981). Structural model indicated that each pillar loads positively and significantly on the composite index (PIRWSS), with $R^2 \approx 0.997$ and bootstrap confidence intervals that exclude zero for all paths, consistent with best-practice SEM reporting (Joseph F Hair et al., 2017). PLS–SEM is appropriate under relaxed distributional assumptions. To enhance operational usability, coefficients were subsequently calibrated via GRG subject to non-negativity, unit-sum, and monotonicity constraints, minimizing prediction error while enforcing interpretability.

Beyond reporting thresholds, the pattern in Table 3 is informative. Constructs such as Documentation (CR = 0.903; AVE = 0.823) and Raw-Water Infrastructure Assets (CR = 0.867; AVE = 0.766) exhibit strong convergent validity, consistent with their highly observable, procedure-bound indicators. By contrast, Organization & Personnel shows a lower α = 0.650 despite acceptable CR = 0.851; this attenuation is expected with few, heterogeneous items and reflects genuine dispersion in managerial capacity across sites rather than measurement weakness. At the second-order level, all three pillars exceed CR ≥ 0.94 and AVE ≥ 0.63, indicating that the latent structure is well captured and justifying use of pillar scores in the calibrated index. High R^2 (≈ 0.997) alone could raise overfitting concerns; however, our bootstrapped paths, cross-validated redundancy (Q^2), and later calibration-stage error checks collectively support predictive rather than merely descriptive fit (Joseph F Hair et al., 2017).

Table 3. Results average variance extracted (AVE)

Dimension	Cronbach's Alpha	α	CR	AVE
A1 = Water Source Quantity	0.917	0.919	0.936	0.710
A2 = Physical Condition of Raw Water Infrastructure	0.858	0.864	0.904	0.702
B1 = Socioeconomic and Cultural	0.930	0.931	0.945	0.741
B2 = Policies/Regulations	0.869	0.871	0.910	0.718
B3 = Organization & Personnel	0.650	0.654	0.851	0.741
B4 = Managing Institution	0.814	0.818	0.890	0.730
B5 = Documentation	0.785	0.787	0.903	0.823
B6 = Raw Water Infrastructure Assets	0.695	0.699	0.867	0.766
B7 = Human Resources	0.759	0.764	0.862	0.675
C1 = Surrounding Environment of Water Source	0.910	0.910	0.930	0.689
C2 = Sustainability of Water Source	0.757	0.758	0.891	0.804
Environmental Variable	0.932	0.932	0.944	0.677
Non-Technical Variable	0.972	0.973	0.974	0.635
Technical Variable	0.944	0.946	0.952	0.667
Performance Index	0.973	0.973	0.975	0.648

3.2. Calibrated Index and Predictive Performance

Following SEM estimation, GRG optimization refined the coefficients under normalization, non-negativity, and monotonicity constraints (Lasdon et al., 1974, 1978). The final closed-form model using equiation (1):

$$PIRWSS = 0.440\ PI_{Tk} + 0.340\ PI_{NT} + 0.220\ PI_{Li} \tag{1}$$

Goodness-of-fit is strong on the composite-index scale, RMSE ≈ 0.522; on the field-score scale, SSR ≈ 83.412 and MSE ≈ 5.721. Because RMSE and MSE are reported on different scales in the dataset (index vs. field scores), they are not directly comparable; the manuscript therefore reports $R^2 \approx 0.997$ (explained variance) for interpretability alongside scale-consistent error metrics. The calibrated hierarchy implies material contributions from all three pillars, with technical drivers exerting the largest effect, followed by non-technical and environmental factors (Figure 3). This ordering coheres with utility-performance literature emphasizing asset condition, capacity, and O&M as first-order determinants, while also recognizing governance enablers and catchment conditions that shape reliability and treatment effort.

The calibrated weights, 0.440 (PITk) > 0.340 (PINT) > 0.220 (PILi), quantify an intuitive hierarchy: immediate performance is most sensitive to asset integrity, continuity, and losses (technical), but it is amplified or constrained by governance (non-technical) and stabilized over the medium term by source-area condition (environmental). Error diagnostics show small site-level deviations between modeled and field scores (absolute differences typically ≤ ~6 points), with the largest positive gap at Otak Aik (field 67.8 vs model 73.6), explained by an exceptional environmental score (10.00) that the field composite underweights and one of the larger negative gaps at Labangka (59.1 vs 55.08), where lower environmental support (7.65) and pumping exposure likely reduce realized performance. Across sites, most absolute deviations are ≤ ~6 points, and sign/magnitude of the gap aligns with pillar imbalances (e.g., very high environmental scores can lift modeled PI above the field composite when governance is only moderate). Interpreting RMSE/MSE alongside R^2 avoids scale confusion (index vs field units) and confirms that calibration improves decision-useful ranking without sacrificing interpretability (Lasdon et al., 1974, 1978). Figure 3 shows clustered bars Lombok sites generally right-shifted (higher PI) due to gravity conveyance and steadier yield, whereas Sumbawa sites left-shift under pumping head and seasonal deficit; this pattern matches rainfall and conveyance contrasts documented for the study area.

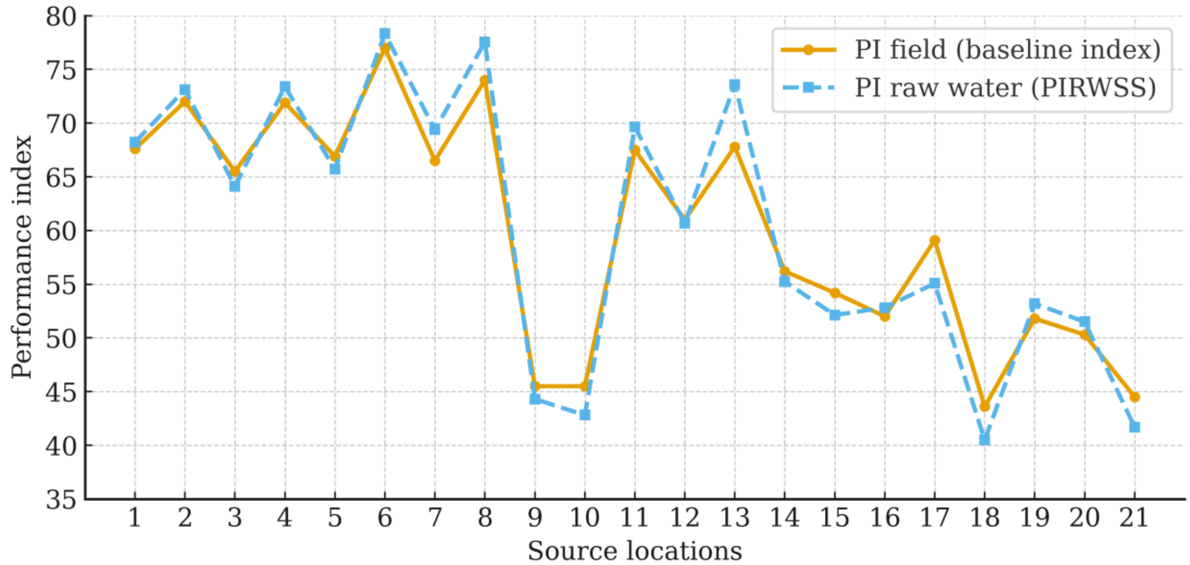

Figure 3. Comparison between field performance index (PI field) and calibrated raw-water performance index (PI raw water) across 21 schemes

3.3. Dimension Level Results

Technical (PI_{Tk}) is indicators of source quantity/reliability and physical asset condition are the strongest drivers of PIRWSS. Gravity-fed schemes in Lombok generally scored higher on reliability and O&M adequacy than pump-dependent systems in Sumbawa, where seasonality and head requirements depress availability and heighten mechanical vulnerabilities. These signals align with resilience and reliability oriented metrics in water-system design and with multi-criteria planning that elevates continuity and condition among top priorities. Non-technical (PI_{NT}) is governance and managerial readiness budgeting, staffing, planning, documentation, and user associations shows a positive, significant path to PIRWSS, confirming that institutions and processes are co-determinants of service outcomes; notably, asset-management and documentation quality emerge as salient levers, consistent with the high loading for the service data book indicator. Environmental (PI_{Li}) is land-cover composition in source areas and source-sustainability proxies provide additional explanatory power and improve alignment with field performance. The pillar's positive, significant path weight indicates that better watershed condition is associated with higher raw-water performance, consistent with evidence that degraded catchments elevate turbidity and nutrients, increase treatment costs, and threaten reliability. Environmental indicators discriminate performance across islands Lombok's more reliable sources and gravity systems support steadier PIRWSS, while Sumbawa's spatially uneven sources and pumping reliance increase vulnerability to demand fluctuations and non-revenue water consistent with the 0.220 composite weight in the integrated index. Performance index shown in Table 4.

A closer examination clarifies why sites diverge despite compliance. Upper-quartile performers (PI ≈ 73–78; Rangat, Sekeper, Serepak, Sarasuta, Otak Aik) pair strong technical scores (15–17) with supportive environmental conditions (8.6–10.0), aided by gravity-fed conveyance and steadier hydrology. Mid-tier sites (60–70) possess adequate assets but are constrained by middling documentation/asset management or moderate source stress, implying tractable O&M remedies. Low performers (<55; Mongglenggo, Rababaka, Sordang, Tibu Ulik) combine weak technical scores (5–10) with pumping head and fragile catchments (7.2–8.1), elevating outages and non-revenue water. These patterns mirror NTB's context: Lombok ≈1,441 mm/year versus Sumbawa ≈1,176 mm/year, and idle capacity (50 L/s designed, 30 L/s delivered).

Table 4 is the technical column is the leading indicator of final PI; values ≥ 15 generally coincide with PI ≥ 70 (e.g., Rangat, Sekeper, Sarasuta). Where non-technical dips < 3.7, scores can be pulled down despite adequate assets (e.g., Pandangduri). Conversely, environmental ≥ 9.5 can offset moderate governance (e.g., Otak Aik). This explains why some sites remain mid-tier although they satisfy individual standards, because composite performance depends on co-movement across pillars rather than any single checklist item.

Table 4. Field performance index and raw water performance index values

Source Location	Technical	Non-Technical	Environmental	Field PI	Raw Water PI
Lebah Sempage	13.0	4.75	9.06	67.6	68.22
Sarasuta	15.0	4.64	9.06	72.0	73.13
Remening	12.0	4.19	9.06	65.5	64.11
Serepak	15.6	4.16	9.06	71.9	73.39
Sesera	12.8	4.46	8.59	66.9	65.72
Rangat	17.0	4.64	9.06	77.0	78.33
Pandangduri	15.0	3.70	8.59	66.5	69.39
Sekeper	17.2	4.60	8.59	74.0	77.54
Tibu Ulik	6.2	3.33	8.12	45.5	44.31
Sordang	5.0	3.93	8.12	45.5	42.81
Singang Pitu Nai	15.0	4.24	8.12	67.5	69.65
Jonplanka	11.6	4.19	8.12	61.0	60.67

Source Location	Technical	Non-Technical	Environmental	Field PI	Raw Water PI
Otak Aik	15.0	3.93	10.00	67.8	73.60
Semongkat	10.0	3.72	8.12	56.2	55.24
Brangdalap	8.8	3.72	8.12	54.2	52.13
Tiu Pasai	9.2	3.59	8.12	52.0	52.82
Labangka	10.2	3.91	7.65	59.1	55.08
Mongglenggo	5.0	3.50	7.65	43.6	40.49
Ncoha	10.4	3.46	7.18	51.8	53.19
Patula	9.8	3.42	7.18	50.3	51.50
Rababaka	6.2	3.25	7.18	44.5	41.69

4. Conclusions

This study set out to construct a decision-ready composite index for raw-water systems and to test it against field reality in NTB. First, a baseline appraisal anchored in SE 03/2021 revealed wide performance dispersion across the 21 schemes, motivating an integrated measure rather than checklist compliance alone. Second, PLS-SEM established that the technical (PI_{Tk}), non-technical (PI_{NT}), and environmental (PI_{Li}) pillars load positively and significantly on overall performance, with strong explanatory power ($R^2 \approx 0.997$). Third, constrained GRG calibration produced an interpretable closed form $PIRWSS = 0.440\ PI_{Tk} + 0.340\ PI_{NT} + 0.220 \cdot PI_{Li}$ that respects non-negativity, unit-sum, and monotonicity. Fourth, validation against field scores showed high agreement (index-scale RMSE ≈ 0.522; field-scale MSE ≈ 5.721, SSR ≈ 83.412), with small site-level deviations explained by pillar imbalances. Finally, the index functions as a transparent, scheme-level benchmark: it prioritizes O&M and rehabilitation where technical deficits dominate, signals managerial reforms when non-technical capacity constrains delivery, and justifies source-area protection when environmental conditions drive stability. These results directly answer the research objectives and deliver a validated, calibrated, and operationally useful index for raw-water performance.

Acknowledgement

We thank the Ministry of Public Works and Public Housing (PWPH) and the Water Resources Public Works Agency (WRPWA) of West Nusa Tenggara Province, Indonesia (NTB) for providing the data. We also appreciate the collaborative efforts of all partners, researchers, and stakeholders in advancing hydrological understanding and in designing management plans that incorporate environmental and ecological considerations.

Ethics Statement

This study involved field observations and questionnaire surveys with adult professional staff of raw-water utilities, basin agencies, and water users in West Nusa Tenggara Province (NTB), Indonesia. According to the applicable institutional and national regulations at the time of data collection, this type of non-clinical, minimal-risk research using anonymized responses from adult professionals did not require review by a formal ethics committee. Participation was voluntary; all respondents were informed about the study objectives, assured of anonymity and confidentiality, and provided their informed consent before completing the questionnaire. No clinical interventions or vulnerable populations were involved.

CRediT Author Statement

Ussy Andawayanti contributed to conceptualization, methodology, formal analysis, writing the original paper, writing the review, editing, and supervision. Ery Suhartanto contributed to conceptualization, methodology, validation, writing review and editing, and supervision. Rahmah Dara Lufira contributed to formal analysis, validation, visualization, and writing review and editing. Hari Siswoyo contributed to supervision, writing review, and editing. Sri Utami Sudiarti contributed to the investigation, data curation, and writing the original paper. Rizki Ramadhani Pratama contributed to the investigation, data curation, and writing the original paper. All authors have read and approved the final version of this manuscript.

References

Alismaiel, O.A., 2021. Using structural equation modeling to assess online learning systems' educational sustainability for university students. Sustainability, 13(24), 13565.

Ayadi, H., Benaissa, M., Hamani, N. and Kermad, L., 2024. Assessing the sustainability of transport systems through indexes: a state-of-the-art review. Sustainability, 16(4), 1455.

Budianto, B., Feri, Z.O., Nurlaila, Q., Suryatman, T.H. and Sitorus, H., 2025. Leagility competencies as mediator hybrid lean agile and operational performance. Vikalpa: The Journal for Decision Makers, 50(3), pp.211–230.

Chakraborty, L., Thistlethwaite, J., Scott, D., Henstra, D., Minano, A. and Rus, H., 2023. Assessing social vulnerability and identifying spatial hotspots of flood risk to inform socially just flood management policy. Risk Analysis, 43(5), 1058–1078.

van Engelenburg, J., van Slobbe, E., Teuling, A.J., Uijlenhoet, R. and Hellegers, P., 2021. Sustainability characteristics of drinking water supply. Drinking Water Engineering and Science, 14(1), pp.1–43.

Fornell, C. and Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), pp.39–50.

Ganjidoost, A., Knight, M.A., Unger, A.J.A. and Haas, C.T., 2021. Performance modeling and simulation for water distribution networks. Frontiers in Water, 3, 718215.

Guo, M., Nowakowska-Grunt, J., Gorbanyov, V. and Egorova, M., 2020. Green technology and sustainable development: assessment and green growth frameworks. Sustainability, 12(16), 6571.

Hair, J.F., Hollingsworth, C.L., Randolph, A.B. and Chong, A.Y.L., 2017a. An updated and expanded assessment of PLS-SEM in information systems research. Industrial Management & Data Systems, 117(3), pp.442–458.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M., 2017b. A primer on partial least squares structural equation modeling (PLS-SEM). 2nd edn. Thousand Oaks, CA: Sage.

He, B., Zheng, H. and Guan, Q., 2024. Toward revolutionizing water–energy–food nexus composite index model: from availability, accessibility and governance. Frontiers in Water, 6, 1338534.

Lasdon, L.S., Fox, R.L. and Ratner, M.W., 1974. Nonlinear optimization using the generalized reduced gradient method. RAIRO – Operations Research, 8(V3), pp.73–103.

Lasdon, L.S., Waren, A.D., Jain, A. and Ratner, M.W., 1978. Design and testing of a generalized reduced gradient code for nonlinear programming. ACM Transactions on Mathematical Software, 4(1), pp.34–50.

Macharia, P., Wirth, M., Yillia, P.T. and Kreuzinger, N., 2021. Examining the relative impact of drivers on energy input for municipal water supply in Africa. Sustainability, 13(15), 8480.

Mao, F., Zhao, X., Ma, P., Chi, S., Richards, K., Clark, J., Hannah, D.M. and Krause, S., 2019. Developing composite indicators for ecological water quality assessment based on network interactions and expert judgment. Environmental Modelling & Software, 115, pp.51–62.

Mergoni, A., Di Inverno, G. and Carosi, L., 2022. A composite indicator for measuring the environmental performance of Portuguese water and waste utilities. Utilities Policy, 74, 101285.

Moudi, M., 2022. Adaptive optimal sustainability framework in urban water supply system under different runoff scenarios. Research Square preprint.

Sankar, J.P., 2024. Spillover effect of workplace politics on work–family conflict: a mediated moderating model. Research Square preprint.

Simpson, G.B., Jewitt, G.P.W., Becker, W., Badenhorst, J., Masia, S., Neves, A.R., Rovira, P. and Pascual, V., 2022. The water–energy–food nexus index: a tool to support integrated resource planning, management and security. Frontiers in Water, 4, 825854.

Su, L. and Cao, Y., 2022. Performance monitoring and evaluation of water environment treatment PPP projects with multi-source heterogeneous information. Frontiers in Environmental Science, 10, 1024701.

Tefera, C.A. and Hunsaker, W.D., 2021. Measurement of intangible assets using higher-order construct model. Journal of Entrepreneurship in Emerging Economies, 13(5), pp.1131–1155.

Zakiyayasin Nisa', S.Q., Novembrianto, R., Ayu Murti, R.H. and Salam Jawwad, M.A., 2023. Sustainability assessment of rural water supply system in Lamongan, Indonesia. International Journal of Eco-Innovation in Science and Engineering, 4(1), 114.