

**RANDOMIZED NUMERICAL LINEAR ALGEBRA FOR THE COMPUTATIONAL
DETECTION OF FAILURES IN THE EXTERNAL EVALUATION OF ANGOLA'S HIGHER
EDUCATION INSTITUTIONS: SELF-AUTHORED PYTHON ALGORITHMS AND AN
ORIGINAL EXPERIMENTAL METHOD IN SUPPORT OF INAAREES**

**ÁLGEBRA LINEAR NUMÉRICA ALEATORIZADA PARA A DETECÇÃO COMPUTACIONAL
DE FALHAS NA AVALIAÇÃO EXTERNA DE INSTITUIÇÕES DE ENSINO SUPERIOR DE
ANGOLA: ALGORITMOS EM PYTHON DE AUTORIA PRÓPRIA E UM MÉTODO
EXPERIMENTAL ORIGINAL EM APOIO AO INAAREES**

 <https://doi.org/10.63330/sasciencesv6n2-226>

Submitted on: 09/26/2026 and Published in: 10/06/2026

SAS: e26453

Pedro Silva

Universidade Gregório Semedo (UGS), Luanda, Angola
ACITE Academia de Ciências Sociais e Tecnologias
Universidade Lusíada de Angola (ULA), Luanda, Angola
E-mail: pedro.silva@ugs.edu.ao
Lattes: <http://lattes.cnpq.br/0794001378193198>
ORCID: <https://orcid.org/0009-0007-0301-5098>

Gerson Jungo

Luanda - Angola, 2026

Abstract

Africa's educational challenge has shifted from quantity to an acute crisis of quality. In Angola, the National Institute for Assessment, Accreditation and Recognition of Higher Education Studies (INAAREES)¹ carried out, between 2023 and 2026, five external evaluation rounds covering 671 study programmes, of which 253 (37.7%) failed accreditation empirical evidence of the subsystem's quality deficit. However, the growing volume of indicators collected in each evaluation cycle makes purely manual analysis slow, costly and error-prone. This paper proposes, for the first time in Angola, the application of randomized numerical linear algebra (RandNLA)² to the computational detection of failures in higher education external evaluations. We introduce an original matrix model of external

¹INAAREES - Instituto Nacional de Avaliação, Acreditação e Reconhecimento de Estudos do Ensino Superior (National Institute for Assessment, Accreditation and Recognition of Higher Education Studies, Angola).

²RandNLA - Randomized Numerical Linear Algebra: a field of computational mathematics that uses random projections to approximate matrix decompositions with probabilistic guarantees.



evaluation (programmes $\times$ indicators), four self-authored Python algorithms SRVD, RSDF, DICA-R and a Failure Risk Index (IRF) and a self-designed experimental method using a synthetic 671×14 matrix calibrated with official INAAREES aggregates, with 60 a priori known injected failures. Results show that SRVD matches the Eckart-Young optimum (maximum deviation of 0.003 percentage points) with speedups up to $37.9\times$ on large matrices; the hybrid RSDF+DICA-R detector recovered 60/60 failures at the top-60 cut-off (100% precision; AUC-PR = 0.976); and the IRF correlates at $r = 0.633$ with the true institutional failure rate. A dedicated analysis of engineering programmes the field assessed in the STEM (4th) phase evidences the robust strand of the national quality assurance framework. Randomization thus provides a viable, scalable and auditable triage instrument for external evaluation committees.

Keywords: Angola, Anomaly detection, Engineering education, External evaluation, Higher education, INAAREES, Python, Quality assurance, Randomized numerical linear algebra.

Resumo

O desafio educacional da África deslocou-se da questão quantitativa para uma crise aguda de qualidade. Em Angola, o Instituto Nacional de Avaliação, Acreditação e Reconhecimento de Estudos do Ensino Superior (INAAREES) realizou, entre 2023 e 2026, cinco ciclos de avaliação externa abrangendo 671 programas de estudo; destes, 253 (37,7%) não obtiveram acreditação — uma evidência empírica do déficit de qualidade do subsistema. No entanto, o crescente volume de indicadores coletados em cada ciclo de avaliação torna a análise puramente manual lenta, dispendiosa e sujeita a erros. Este artigo propõe, pela primeira vez em Angola, a aplicação da álgebra linear numérica aleatorizada (RandNLA) para a detecção computacional de falhas em avaliações externas do ensino superior. Apresentamos um modelo matricial original de avaliação externa (programas $\times$ indicadores), quatro algoritmos em Python desenvolvidos pelos autores — SRVD, RSDF, DICA-R e um Índice de Risco de Falha (IRF) — e um método experimental próprio utilizando uma matriz sintética de 671×14 , calibrada com dados agregados oficiais do INAAREES e contendo 60 falhas inseridas artificialmente (conhecidas *a priori*). Os resultados



demonstram que o SRVD iguala o ótimo de Eckart-Young (desvio máximo de 0,003 pontos percentuais) com acelerações de até 37,9 vezes em matrizes grandes; o detector híbrido RSDF+DICA-R identificou todas as 60 falhas no corte dos 60 primeiros resultados (precisão de 100%; AUC-PR = 0,976); e o IRF apresenta uma correlação de $r = 0,633$ com a taxa real de falha institucional. Uma análise específica dos programas de engenharia — a área avaliada na fase STEM (4ª fase) — evidencia a robustez do sistema nacional de garantia da qualidade. A aleatorização oferece, assim, um instrumento de triagem viável, escalável e auditável para as comissões de avaliação externa.

Palavras-chave: Álgebra linear numérica aleatorizada, Angola, Avaliação externa, Detecção de anomalias, Ensino de engenharia, Ensino superior, Garantia da qualidade, INAAREES, Python.

INTRODUCTION

The debate on education in Africa has undergone a decisive conceptual shift over the past two decades: while the historical priority was expanding access enrolment, infrastructure, institutional networks — recent literature and public policies converge on the recognition that the continent's central problem is now a crisis of quality, not of quantity [18]. In higher education this shift is particularly visible: the multiplication of institutions and programme offerings has not always been matched by the technical-pedagogical and scientific conditions that confer academic and social relevance upon them [10, 11].

In Angola, this concern is institutionalized in INAAREES, a public institute endowed with legal personality and administrative autonomy, whose Organic Statute was approved by Presidential Decree No. 172/13 of 29 October [14]. Its legal mission is to “promote and monitor the quality of the technical-pedagogical and scientific conditions created and of the services provided by higher education institutions” [14], framed by the Legal Regime for the Evaluation and Accreditation of the Quality of Higher Education Institutions (Presidential Decree No. 203/18 of 30 August) [12], by the Legal Regime of the Higher Education Subsystem (Presidential Decree No. 310/20 of 7 December) [13] and by the



Regulation of the External Evaluation and Accreditation Process (Executive Decree No. 109/20 of 10 March, amended by Executive Decree No. 148/24) [15].

The aggregate results of the first five phases of the national external evaluation process, released by INAAREES, confirm the severity of the quality deficit: of the 671 programmes evaluated between October 2023 and 2026 medicine and health sciences programmes (1st phase and re-evaluation), the remaining health offerings (2nd phase), education sciences (3rd phase), STEM³ programmes (4th phase) and humanities, arts and social sciences programmes (HASS⁴) (5th phase) only 418 obtained accreditation or conditional accreditation, while 253 (37.7%) were not accredited [16]. In the 4th phase, dedicated to STEM fields, 445 specialists were mobilized 356 external evaluators and 89 procedure managers, including experts from Brazil, Cape Verde, Cuba, Mozambique, the Democratic Republic of the Congo, Zimbabwe, Portugal and São Tomé and Príncipe to evaluate 241 programmes from 70 institutions [17]. At the continental level, the HAQAA⁵ initiative assigned Angola a score of 3.6 out of 5 for compliance with international quality assurance standards, placing the country below the African average [17].

Each evaluation cycle generates a growing mass of data: dozens of indicators per programme, hundreds of programmes per phase, multiple sources of evidence (documentary, interviews, site visits). The analysis of these data remains, however, essentially manual and aggregative means, percentages and classification tables which limits the capacity to (i) detect severe, localized failures diluted in acceptable averages, (ii) identify internal inconsistencies between indicator families (for example, declared qualified teaching staff without corresponding scientific output) and (iii) rationally prioritize the institutions and programmes that should receive reinforced scrutiny from the external evaluation committees.

It is precisely in this gap that the present article is situated. Randomized numerical linear algebra (RandNLA) is a consolidated field of computational mathematics that uses random projections to

³STEM - Science, Technology, Engineering and Mathematics.

⁴HASS - Humanities, Arts and Social Sciences.

⁵HAQAA - Harmonisation of African Higher Education Quality Assurance and Accreditation (African initiative for the harmonisation of quality assurance and accreditation in higher education).



approximate, with rigorous probabilistic guarantees, classical objects such as the singular value decomposition (SVD)⁶, with order-of-magnitude scalability gains [1, 5]. However, to the best of the author's knowledge, no published study has applied RandNLA techniques to external evaluation data in Angolan higher education nor, more generally, to African quality assurance systems. This article aims to fill that gap, standing as the first methodological contribution of its kind in the country, developed within the Universidade Gregório Semedo, the Universidade Metodista de Angola and the Universidade Lusíada de Angola.

Contributions. The article offers five original contributions:

- (i) a matrix model of external evaluation (Section 3): formalization of the evaluation matrix (programmes $\times$ indicators) and empirical demonstration of its effectively low-rank structure (99.78% of spectral energy concentrated in four factors);
- (ii) three self-authored Python algorithms (Section 4): SRVD⁷ (randomized SVD with oversampling and power iteration), RSDF⁸ (randomized spectral residual for failure detection) and DICA-R⁹ (detection of incoherence by random compression with repeated random partitions of the indicators);
- (iii) a Failure Risk Index (IRF)¹⁰ (Section 4.4): a composite institutional prioritization metric combining spectral anomaly and non-compliance with the legal accreditation threshold (60%);
- (iv) a dedicated analysis of engineering programmes (Section 6.1): a sectoral reading of the robust strand that engineering and technology programmes bring to the national quality assurance landscape;

⁶SVD - Singular Value Decomposition.

⁷SRVD - Randomized SVD with oversampling (self-authored algorithm, Algorithm 1 of this paper).

⁸RSDF - Randomized Spectral Residual for Failure Detection (self-authored algorithm, Algorithm 2).

⁹DICA-R - Detection of Incoherence by Randomized Random Compression (self-authored algorithm, Algorithm 3).

¹⁰IRF - Failure Risk Index (composite institutional prioritization metric, Algorithm 4).



(v) a self-designed experimental method (Sections 3 and 5): a simulation protocol calibrated with official INAAREES aggregates (671 programmes; accreditation rate $\approx 62\%$; 60 injected failures with known ground truth), allowing the detection capacity to be measured objectively (AUC-ROC¹¹, AUC-PR¹², precision at the operational cut-off) before any pilot with real data.

Article structure. Section 2 presents the theoretical background; Section 3 describes materials and methods; Section 4 details the proposed algorithms; Section 5 reports the experimental results; Section 6 discusses implications including the engineering strand, limitations and ethical considerations; Section 7 concludes.

THEORETICAL BACKGROUND

THE SINGULAR VALUE DECOMPOSITION AND THE ECKART–YOUNG THEOREM

Let $A \in \mathbb{R}^{m \times n}$ be the evaluation matrix, with m programmes and n indicators. The singular value decomposition (SVD) is written as:

$$A = U \Sigma V^T = \sum_j \sigma_j u_j v_j^T, \quad \sigma_1 \geq \sigma_2 \geq \dots \geq 0 \quad (1)$$

where U and V are orthonormal and σ_j are the singular values. The Eckart–Young theorem [7] guarantees that the best rank- k approximation, in Frobenius norm, is obtained by truncation:

$$\min_{\text{rank}(B) \leq k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sum_{j > k} \sigma_j^2} \quad (2)$$

The working hypothesis of this article verified empirically in Section 5.1 is that external evaluation matrices are dominated by a few latent quality profiles (robust, consolidating, fragile and

¹¹AUC-ROC - Area Under the Receiver Operating Characteristic curve.

¹²AUC-PR - Area Under the Precision–Recall curve.



critical institutions); that is, spectral energy concentrates in the first singular values and the spectral tail decays rapidly. It is precisely this regime that makes randomized methods effective [1].

THE JOHNSON-LINDENSTRAUSS LEMMA AND RANDOM PROJECTIONS

The Johnson–Lindenstrauss lemma [8] states that m points in $\mathbb{R}^n$ can be projected into a space of dimension $d = O(\varepsilon^{-2} \log m)$ while preserving all Euclidean distances up to a distortion ε :

$$(1 - \varepsilon)\|x_i - x_j\|^2 \leq \|\Omega^T x_i - \Omega^T x_j\|^2 \leq (1 + \varepsilon)\|x_i - x_j\|^2 \quad (3)$$

where $\Omega \in \mathbb{R}^{n \times d}$ is a Gaussian random matrix scaled by $1/\sqrt{d}$. The practical consequence is immediate: the proximity geometry between programmes — essential for detecting “atypical” programmes — survives aggressive random compression, which drastically reduces computational cost [3].

RANDOMIZED SVD

The canonical algorithm of Halko, Martinsson and Tropp [1] approximates the dominant column space of A by sampling it with Gaussian vectors: one computes $Y = A\Omega$, with $\Omega \in \mathbb{R}^{n \times (k+p)}$, obtains an orthonormal basis Q of Y by QR factorization, projects $B = Q^T A$ and factorizes the small matrix B . The oversampling parameter p (typically 5–10) and the power iteration $Y \leftarrow A(A^T Y)$, repeated q times, control the quality of the approximation. The expected probabilistic guarantee is [1]:

$$E\|A - QQ^T A\|_F \leq \sqrt{1 + \frac{k}{p-1}} \cdot \sqrt{\sum_{j>k} \sigma_j^2}. \quad (4)$$

In other words, the randomized error lies within a factor $\sqrt{1 + k/(p-1)}$ of the Eckart-Young optimum — for $k = 4$ and $p = 6$, a theoretical factor of ≈ 1.34 which in practice proves much tighter



(Section 5.1). Comprehensive reviews of the field can be found in Mahoney [2], Woodruff [3], Drineas and Mahoney [4] and Martinsson and Tropp [5]; practical sketching algorithms for low-rank approximation are discussed by Tropp et al. [6].

ANOMALY DETECTION AND ROBUST STANDARDIZATION

A programme “with failure” is, in the language of this article, a programme whose row of the evaluation matrix departs from the subspace that describes the system's regular behaviour — either because some indicators drop abruptly (severe non-conformity), or because the declared indicators are mutually inconsistent. To turn that departure into a decision, we use robust standardization based on the median and the median absolute deviation (MAD)¹³, insensitive to extreme values:

$$z_i = \frac{s_i - \text{median}(s)}{1.4826 \cdot \text{MAD}(s)} \quad (5)$$

The factor 1.4826 calibrates the MAD as a consistent estimator of the standard deviation under normality, allowing z_i to be interpreted as a “number of robust deviations” above typical behaviour.

MATERIALS AND METHODS

RESEARCH DESIGN

This is an experimental-computational study, with a self-authored method, organized into five experiments (E1–E5): accuracy of the randomized SVD (E1), computational scalability (E2), distance preservation under random projection (E3), failure detection with controlled ground truth (E4) and institutional prioritization (E5). Since INAAREES primary data, disaggregated by indicator, are not public, a calibrated simulation protocol was adopted: the synthetic matrix reproduces the dimensions and official aggregates of the real process (671 programmes evaluated; overall accreditation rate $\approx 62\%$ [16]),

¹³MAD - Median Absolute Deviation.



and failures are injected artificially and known a priori which makes it possible to measure objectively the precision and recall of the detectors. This design is the standard for methodological validation in RandNLA [1, 6] and necessarily precedes any pilot with real data.

THE MATRIX MODEL OF EXTERNAL EVALUATION

We propose representing each external evaluation cycle by a matrix $A \in \mathbb{R}^{m \times n}$, in which each row corresponds to an evaluated programme and each column to a quality indicator, scored on a scale of 0 to 100 the INAAREES External Evaluation Manual [9] provides for the translation of evidence into scores, with 60% as the accreditation threshold [10]. For this study we defined 14 indicators representative of the dimensions usually evaluated (Table 1), organized into four families: academic capital, material resources, management processes and engagement with society.

Table 1

Indicators of the proposed matrix model (0–100 scale), by quality family.

Code	Indicator	Family	Code	Indicator	Family
I01	Teaching staff qualifications	Academic capital	I08	Internal self-assessment system	Management
I02	Curricular adequacy	Academic capital	I09	Graduate employability	Engagement with society
I03	Physical infrastructure	Material resources	I10	University outreach	Engagement with society
I04	Library and bibliographic resources	Material resources	I11	Information technologies	Material resources
I05	Laboratories and equipment	Material resources	I12	Tutoring and student support	Management
I06	Scientific research	Academic capital	I13	Learning assessment	Academic capital
I07	Academic and administrative management	Management	I14	Cooperation and internationalization	Engagement with society

Source: Original figure by the author (2026).



GENERATION AND CALIBRATION OF THE SYNTHETIC MATRIX

The synthetic matrix was generated by a latent-factor model of rank $r = 4$, interpretable as four quality profiles of the system (robust, consolidating, fragile and critical):

$$A_{clean} = clip((M \odot N)F + \eta), \quad M \in \mathbb{R}^{m \times r}, \quad F \in \mathbb{R}^{r \times n}, \quad \eta \sim N(0; 1.2^2) \quad (6)$$

where M contains each programme's fuzzy membership coefficients to the profiles (Dirichlet distribution), $N = [89.0; 74.5; 60.5; 44.0]$ fixes each profile's mean level and F encodes each profile's signature on the 14 indicators, with a structural 15% penalty on indicators I03, I05 and I14 reflecting the chronic weakness of infrastructure, laboratories and internationalization reported in the national literature [11]. The levels were calibrated until the share of programmes with mean ≥ 60 in the final matrix (65.9%) approached the official aggregate (62.3% accredited in the last three phases [16]). The resulting matrix has 99.78% of spectral energy concentrated in the first four singular values ($\sigma_1 = 6176.8$; $\sigma_5/\sigma_1 \approx 0.021$), confirming the effectively low-rank regime.

FAILURE INJECTION (GROUND TRUTH)

Sixty known failures were injected, under two severity regimes:

- (a) Severe failures (42):** a random reduction of 22-40 points in 2 to 5 indicators of the same programme, concentrated in 14 of the 70 simulated institutions emulating the systemic fragility of problematic institutions;
- (b) Mild failures (18):** a reduction of 10-18 points in 1 to 2 indicators, randomly dispersed — emulating localized non-conformities.

The concentration of severe failures in a few institutions replicates the real pattern of quality assurance systems, in which non-conformities cluster around common institutional factors (governance, funding, shared teaching staff) [10].



ALGORITHM EVALUATION METRICS

Accuracy: relative Frobenius error $\|A - A_k\|/\|A\|$, compared with the Eckart–Young optimum (Eq. 2); 30 independent repetitions per configuration;

Scalability: median runtime of the exact SVD (LAPACK)¹⁴ vs. the randomized one, on matrices from 671×14 up to $20,000 \times 2,000$;

Geometric preservation: maximum distortion of Euclidean distances between 400 programme pairs after random projection of dimension d ;

Detection: AUC-ROC, AUC-PR and precision at the top-60 operational cut-off (equal to the number of injected failures);

Prioritization: Pearson correlation between the institutional IRF and the true rate of injected failures per institution.

All computations were executed in Python 3.12 with NumPy, with a fixed random seed (2026) for full reproducibility; reported times are medians of repetitions on the same machine. The algorithms, synthetic data, charts and experimental method are the author's own work.¹⁵

PROPOSED ALGORITHMS (SELF-AUTHORED)

Figure 1 summarizes the end-to-end pipeline of the proposed method: the evaluation matrix (programmes $\times$ indicators) is compressed by the randomized SVD (Algorithm 1); two complementary detectors then operate on the compressed representation RSDF measures how far each programme lies from the dominant subspace (Algorithm 2), while DICA-R tests whether each programme's indicators are mutually predictable under repeated random partitions (Algorithm 3); finally, the normalized hybrid score feeds the IRF, which ranks institutions by jointly weighing mean anomaly and legal non-compliance (Algorithm 4). All computations run in seconds on ordinary hardware.

¹⁴LAPACK - Linear Algebra PACKage (reference numerical linear algebra library, used internally by NumPy).

¹⁵All charts, tests, tables and computations presented in this article are the author's own work.

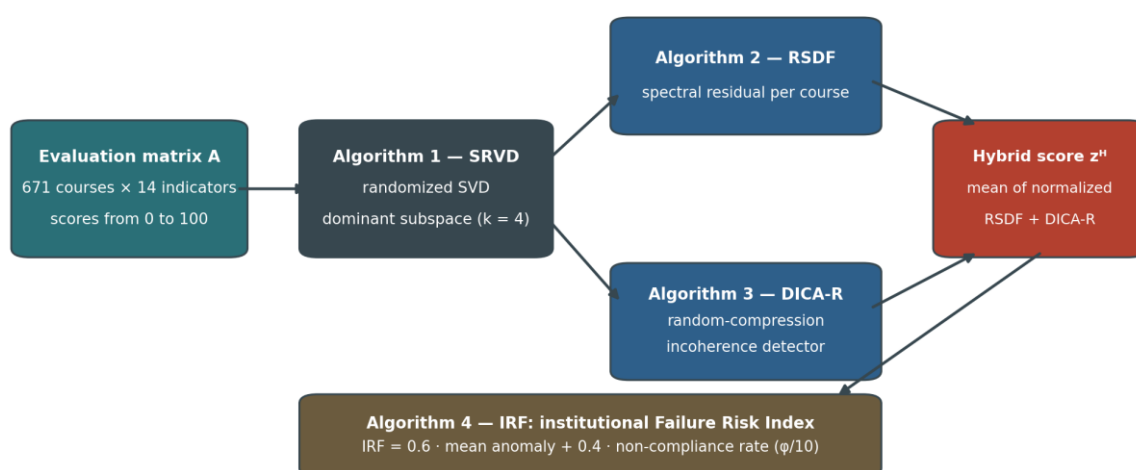


Figure 1

End-to-end computational triage pipeline proposed for INAAAREES external evaluation. Reading guide: the evidence matrix enters on the left; the SRVD compresses it into the dominant quality subspace; the upper branch (RSDF) captures spectral deviations and the lower branch (DICA-R) captures internal inconsistencies; both feed the hybrid score per programme, which the IRF aggregates into an institutional risk ranking that supports committee triage.

Computational triage pipeline proposed for INAAAREES external evaluation

Randomized linear algebra turns a 671×14 evidence matrix into a ranked, auditable risk list in seconds



Source: Original figure by the author (2026).

ALGORITHM 1 - SRVD: RANDOMIZED SVD WITH OVERSAMPLING

The first algorithm implements, from scratch, the Halko–Martinsson–Tropp scheme [1], with oversampling p and q power iterations to refine the spectrum necessary when the spectral tail decays slowly, as happens after failure injection (Section 3.4).



Algorithm 1 - SRVD (self-authored implementation in Python/NumPy)

```
import numpy as np

def randomized_svd(A, k, p=6, q=2, seed=2026):
    """Rank-k randomized SVD with oversampling p
    and q power iterations."""
    rng = np.random.default_rng(seed)
    m, n = A.shape
    Omega = rng.standard_normal((n, k + p)) # 1. Gaussian test matrix
    Y = A @ Omega # 2. sample the column space
    for _ in range(q): # 3. power iteration
        Y = A @ (A.T @ Y) # (amplifies dominant spectrum)
    Q, _ = np.linalg.qr(Y, mode='reduced') # 4. orthonormal basis (m, k+p)
    B = Q.T @ A # 5. low-cost projection (k+p, n)
    Ub, S, Vt = np.linalg.svd(B, full_matrices=False) # 6. SVD of the small matrix
    return (Q @ Ub)[: , :k], S[:k], Vt[:k, :]
```

The dominant complexity is $O(mn(k+p))$ for the multiplications by Ω - against $O(mn \cdot \min(m,n))$ of the exact SVD - which explains the speedup observed in Section 5.2.

ALGORITHM 2 - RSDF: RANDOMIZED SPECTRAL RESIDUAL FOR FAILURE DETECTION

The intuition behind RSDF is geometric: if the system's regular programmes live approximately in a low-dimensional subspace (the latent profiles), then a programme with severe failures will project an anomalous component outside that subspace. RSDF estimates the subspace with SRVD (avoiding the cost of the exact SVD) and measures, for each programme i , the norm of its row residual after orthogonal projection onto the approximated row space:

$$r_i = \|a_i(I - VV^T)\|_2, \quad z_i^{RSDF} = \frac{r_i - \bar{r}}{1.4826 \cdot MAD(r)} \quad (7)$$



Algorithm 2 - RSDF (self-authored implementation)

```
def rsdf_scores(A, k=4, p=6, q=2, seed=2026):
    """Per-course anomaly score: norm of the residual outside the
    dominant subspace approximated by randomized SVD."""
    _, _, Vt = randomized_svd(A, k, p, q, seed)
    residual = A - (A @ Vt.T) @ Vt      # row-wise orthogonal projection
    s = np.linalg.norm(residual, axis=1) # residual norm (m,)
    med = np.median(s)
    mad = np.median(np.abs(s - med))
    return (s - med) / (1.4826 * mad)    # robust z-score per course
```

ALGORITHM 3 - DICA-R: DETECTION OF INCOHERENCE BY RANDOMIZED RANDOM COMPRESSION

DICA-R is, to the author's knowledge, a new algorithm, designed for a type of failure that the spectral residual may miss: the internal inconsistency of declared indicators for example, a programme that declares highly qualified teaching staff (I01) but whose research (I06) and learning assessment (I13) are incompatible with that declaration. The idea is to exploit the dependency structure between indicator families: in a regular system, one half of the indicators predicts the other half through the compressed coordinates of the global subspace. By repeating random partitions of the indicators over T iterations, each cell of the matrix is “tested” against the remaining ones dozens of times; inconsistent programmes accumulate prediction error:

$$W_t = (Z^T Z + \lambda I)^{-1} Z^T A_{:,S_2(t)}, \quad z_i^{DICA} \propto \frac{1}{T} \sum_{t=1..T} \|A_{i,S_2(t)} - z_i^T W_t\|^2 \quad (8)$$

where $Z = A\hat{V}$ are the compressed coordinates obtained by SRVD and $\{S_1(t), S_2(t)\}$ is the random partition of the indicators at iteration t . Ridge regression with $\lambda = 1$ stabilizes the prediction map.



Algorithm 3 - DICA-R (self-authored implementation)

```
def dica_r(A, T=60, k=4, lam=1.0, seed=2026):
    """Detection of incoherence by randomized random compression.
    In each of the T iterations, the indicators are randomly split
    into two halves; a ridge map learns to predict the 2nd half from
    the compressed coordinates; inconsistent courses accumulate
    prediction error."""
    rng = np.random.default_rng(seed)
    m, n = A.shape
    U, Vt = randomized_svd(A, k, p=4, q=2, seed=seed)
    Z = A @ Vt.T * np.sqrt(n / (2 * k)) # compressed coordinates (m, k)
    acum = np.zeros(m)
    for t in range(T):
        perm = rng.permutation(n)
        S2 = perm[n // 2:] # random 'test' half
        W = np.linalg.solve(Z.T @ Z + lam * np.eye(k), Z.T @ A[:, S2])
        E = A[:, S2] - Z @ W # cross-prediction error
        acum += np.linalg.norm(E, axis=1)
    s = acum / T
    med = np.median(s); mad = np.median(np.abs(s - med))
    return (s - med) / (1.4826 * mad)
```

ALGORITHM 4 - HYBRID SCORE AND THE FAILURE RISK INDEX (IRF)

The hybrid score combines the two normalized pieces of evidence, exploiting the complementarity between spectral deviation and internal inconsistency:

$$z_i^H = \frac{1}{2} \left(\frac{z_i^{RSDF}}{\hat{\sigma}_{RSDF}} + \frac{z_i^{DICA}}{\hat{\sigma}_{DICA}} \right) \quad (9)$$

To support institutional decision-making which institutions should receive reinforced committees or additional site visits the scores of each institution *j*'s programmes are aggregated into the Failure Risk Index:



$$IRF_j = 0.6 \cdot \bar{z}_j^+ + 0.4 \cdot \frac{\varphi_j}{10} \quad (10)$$

where $\bar{z}_j^+$ is the mean of the positive scores of institution j 's programmes and φ_j the percentage of its programmes below the legal accreditation threshold of 60 points [9, 10]. The first term measures mean anomaly; the second, exposure to non-compliance. The 0.6/0.4 weights were fixed a priori and not fitted to the data, and should be revisited in a pilot with real data.

EXPERIMENTAL RESULTS

E1 ACCURACY OF THE RANDOMIZED SVD

Table 2 and Figure 2 compare, for target ranks $k = 1, \dots, 12$, the relative Frobenius error of the approximation obtained by SRVD (mean and standard deviation of 30 repetitions) with the Eckart–Young theoretical optimum. The maximum observed deviation was 0.003 percentage points (at $k = 4$: 4.680% vs. 4.677%), with negligible variance randomization does not, in practice, sacrifice accuracy in this effectively low-rank regime.

Table 2

Relative Frobenius error (%) of the rank- k approximation: optimum (exact SVD) vs. SRVD (30 repetitions).

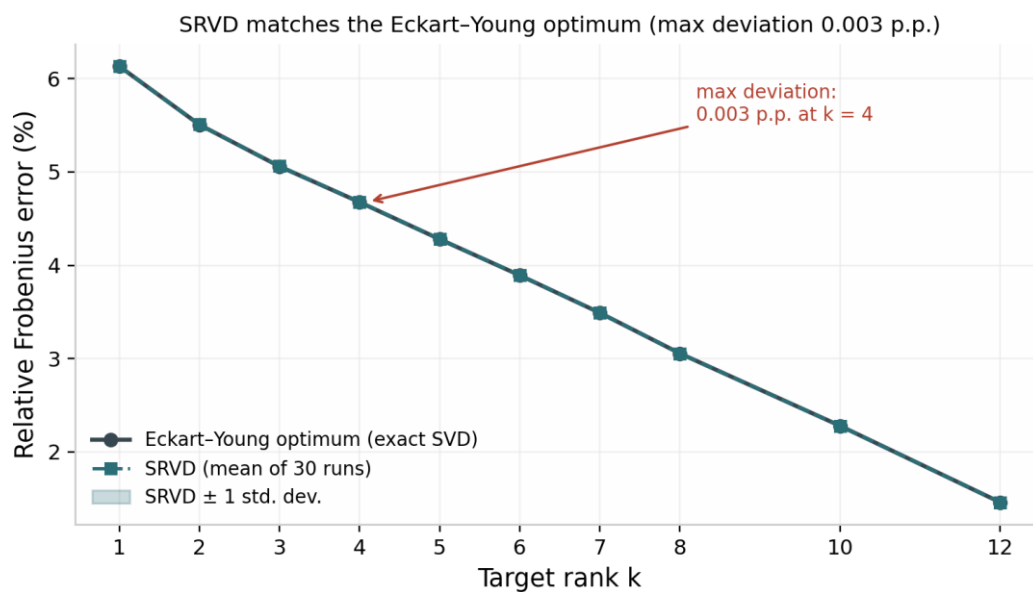
k	Optimum (Eckart–Young)	SRVD (mean)	SRVD (std. dev.)	Deviation (p.p.)
1	6.133	6.133	< 0.001	< 0.001
2	5.507	5.508	0.001	0.001
3	5.060	5.063	0.003	0.003
4	4.677	4.680	0.003	0.003
5	4.280	4.283	0.004	0.003
6	3.893	3.895	0.002	0.002
7	3.491	3.492	0.001	0.001
8	3.056	3.056	< 0.001	< 0.001
10	2.280	2.280	< 0.001	< 0.001
12	1.460	1.460	< 0.001	< 0.001

Source: Original figure by the author (2026).



Figure 2

Accuracy of the randomized SVD (SRVD) against the Eckart-Young optimum on the 671×14 evaluation matrix. How to read: the lower the curve, the better the approximation; the dashed teal curve (mean of 30 independent runs) is visually indistinguishable from the exact optimum (dark solid line), and the shaded band (± 1 standard deviation) is nearly invisible — evidence of the estimator's stability. The largest observed deviation, 0.003 percentage points at $k = 4$, is highlighted by the arrow.



Source: Original figure by the author (2026).

E2 - COMPUTATIONAL SCALABILITY

To gauge behaviour in larger evaluation cycles (more programmes, more indicators, time series across several cycles), we measured the median runtime of the exact SVD (LAPACK, via `numpy.linalg.svd`) and of SRVD with $k = 10$, $p = 6$, $q = 2$ on growing matrices (Table 3, Figure 3). At the current dimension of the Angolan process (671×14) both are instantaneous; but as size grows, the cubic cost of the exact SVD explodes, while SRVD scales almost linearly 37.9× faster at $20,000 \times 2,000$ (24.3 s vs. 0.64 s). This result is decisive for the future exploitation of INAAREES longitudinal databases, combining multiple cycles, programmes and indicators disaggregated by dimension and subdimension.



Table 3

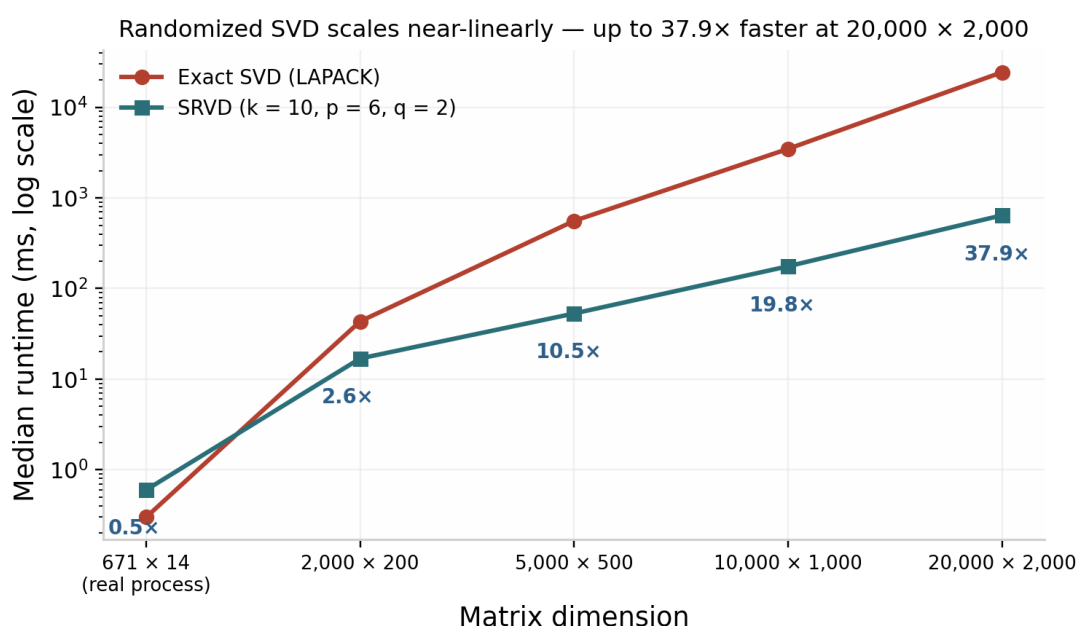
Median runtime of the exact SVD vs. SRVD ($k = 10$, $p = 6$, $q = 2$), in milliseconds.

Matrix dimension	Exact SVD (ms)	SRVD (ms)	Speedup
671×14 (real process)	0.3	0.6	0.5×
$2,000 \times 200$	43.4	16.9	2.6×
$5,000 \times 500$	557.2	53.0	10.5×
$10,000 \times 1,000$	3,473.7	175.8	19.8×
$20,000 \times 2,000$	24,343.3	642.2	37.9×

Source: Original figure by the author (2026).

Figure 3

Scalability of the exact SVD vs. the randomized SVD (logarithmic scale). How to read: the vertical axis is logarithmic, so the widening gap between the red curve (exact SVD) and the teal curve (SRVD) represents an exponentially growing advantage; the blue annotations give the speedup factor at each dimension. At the current size of the Angolan evaluation process (671×14) the difference is irrelevant; the advantage becomes structural for large longitudinal databases.



Source: Original figure by the author (2026).

E3 - DISTANCE PRESERVATION (JOHNSON-LINDENSTRAUSS)

We projected the evaluation matrix into spaces of dimension $d \in \{10, 20, 40, 80\}$ with scaled Gaussian matrices and measured the maximum distortion of Euclidean distances between 400 programme pairs (Table 4, Figure 4). With $d = 80$ the maximum observed distortion was 0.193, empirically



confirming the Johnson–Lindenstrauss lemma on this dataset. The operational reading is relevant for INAAREES: proximity analyses between programmes (clustering, profile comparison, detection of atypical neighbours) can be conducted in a compressed space, with controlled losses.

Table 4

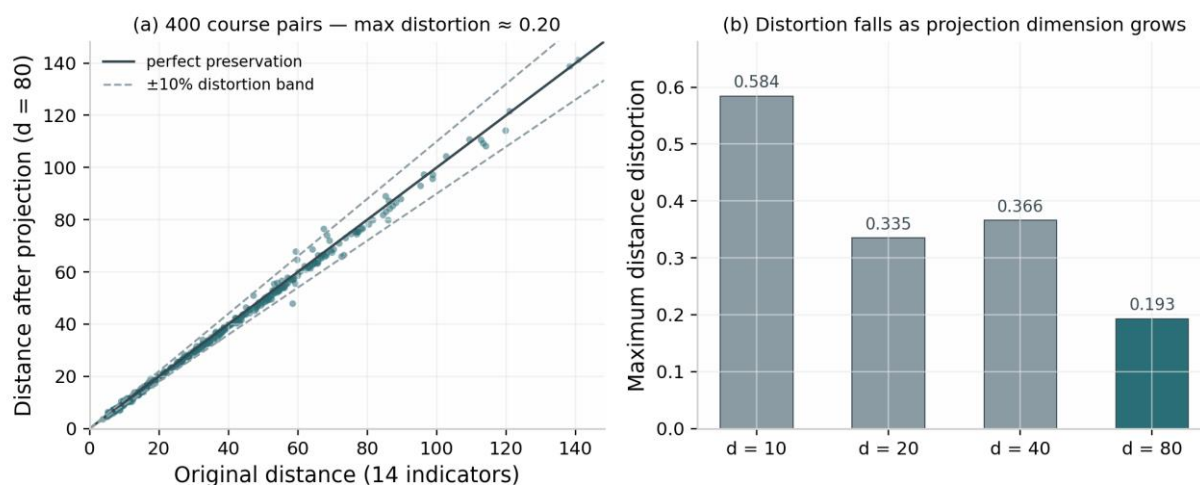
Maximum distortion of distances between 400 programme pairs after random projection.

Projected dimension d	10	20	40	80
Max. distance distortion (400 pairs)	0.584	0.335	0.366	0.193

Note: the slightly higher distortion at $d = 40$ than at $d = 20$ reflects random variability between independent projections, consistent with the probabilistic nature of the lemma. Source: Original figure by the author (2026).

Figure 4

Euclidean distances in the original space (14 indicators) vs. distances after random projection. How to read: in panel (a), each point is one of 400 programme pairs; points lying on the solid line have their distance exactly preserved, and the dashed lines delimit the $\pm 10\%$ distortion band at $d = 80$ nearly all pairs fall inside the band (maximum distortion ≈ 0.19). Panel (b) shows how the maximum distortion (bars) decreases as the projection dimension grows, with $d = 80$ (teal bar) achieving the best fidelity.



Source: Original figure by the author (2026).

E4 - FAILURE DETECTION WITH CONTROLLED GROUND TRUTH

We applied the three detectors (RSDF, DICA-R and the hybrid) to the matrix with the 60 injected failures, without any prior knowledge of their location. The results (Table 5, Figures 5 and 7) are unequivocal: at the top-60 operational cut-off which corresponds to flagging exactly as many



programmes as there are failures all three detectors recovered 60 out of 60 real failures (100% precision), including the 18 mild failures of only 10-18 points in one or two indicators. The precision–recall curves reach $AUC-PR = 0.976$, maintaining unit precision until near-total recall, and $AUC-ROC$ is 1.000 for all three methods.

Table 5

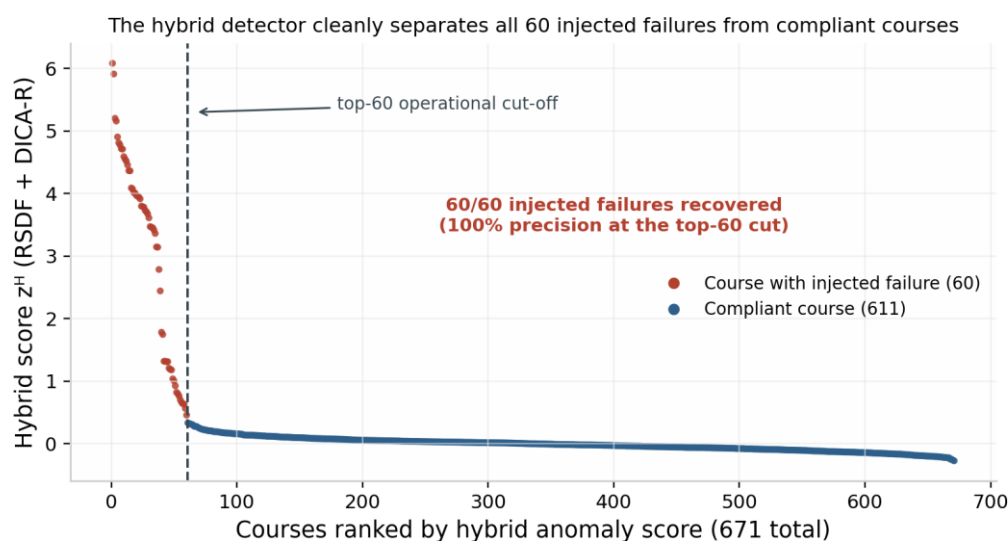
Performance of the failure detectors (60 failures injected into 671 programmes).

Detector	AUC-ROC	AUC-PR	Top-60 precision	Severe failures (42)	Mild failures (18)
RSDF (randomized spectral residual)	1.000	0.976	60/60 (100%)	42/42	18/18
DICA-R (incoherence by random compression)	1.000	0.976	60/60 (100%)	42/42	18/18
Hybrid RSDF + DICA-R	1.000	0.976	60/60 (100%)	42/42	18/18

Source: Original figure by the author (2026).

Figure 5

Hybrid anomaly scores of the 671 programmes, sorted in descending order, from a faithful re-execution of the protocol described in Sections 3.3–3.4. How to read: each point is a programme; red points are programmes with real injected failures and blue points are compliant programmes. The dashed vertical line marks the top-60 operational cut-off flagging exactly as many programmes as there are failures and the sharp drop right at that line shows that the 60 highest scores correspond exactly to the 60 injected failures (100% precision).



Source: Original figure by the author (2026).



E5 - INSTITUTIONAL PRIORITIZATION WITH THE IRF

Aggregating the hybrid scores at the level of the 70 simulated institutions (Eq. 10), the IRF produced a risk ranking whose correlation with the true rate of injected failures per institution was $r = 0.633$ (Table 6, Figure 6). The correlation is not perfect for a substantive and interesting reason: the IRF captures, beyond the injected failures, the structural fragility of institutions institutions without injected failures but with a high percentage of programmes below the 60-point threshold (for example, IES-55 and IES-53, with 75% and 71% of programmes below the threshold) legitimately appear at the top of the ranking. In a real application, it is precisely this dual reading (anomaly + non-compliance) that matters to the regulator.

Table 6

Eight institutions with the highest IRF (simulated data): index composition and true rate of injected failures.

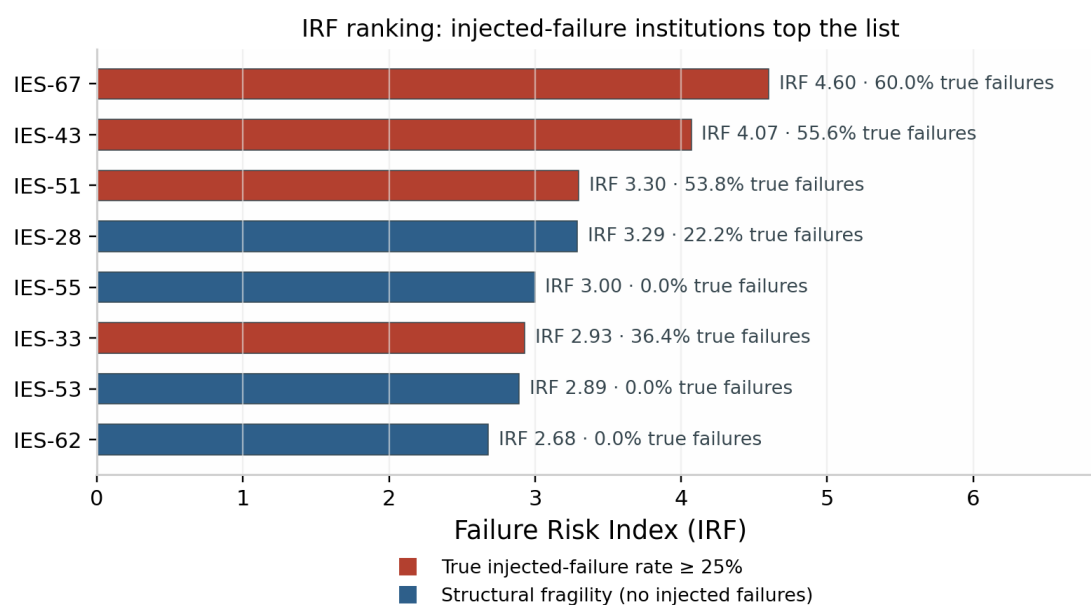
Institution	No. of programmes	% programmes < 60 (φ)	Mean anomaly ($\bar{z}^+$)	IRF	True failure rate
IES-67	10	80.0	2.34	4.60	60.0%
IES-43	9	66.7	2.34	4.07	55.6%
IES-51	13	53.8	1.90	3.30	53.8%
IES-28	9	66.7	1.05	3.29	22.2%
IES-55	8	75.0	0.01	3.00	0.0%
IES-33	11	54.5	1.24	2.93	36.4%
IES-53	7	71.4	0.06	2.89	0.0%
IES-62	9	66.7	0.02	2.68	0.0%

Source: Original figure by the author (2026).



Figure 6

Ranking of the institutions with the highest IRF (simulated data). How to read: bars show the composite Failure Risk Index; red bars identify institutions whose true injected-failure rate is $\geq 25\%$, and the label at the end of each bar reports the IRF value and the true failure rate. The blue bars near the top (IES-28, IES-55, IES-53, IES-62) correspond to structural fragility many programmes below the 60-point threshold without injected point failures which is precisely the kind of risk the regulator wants to see flagged.

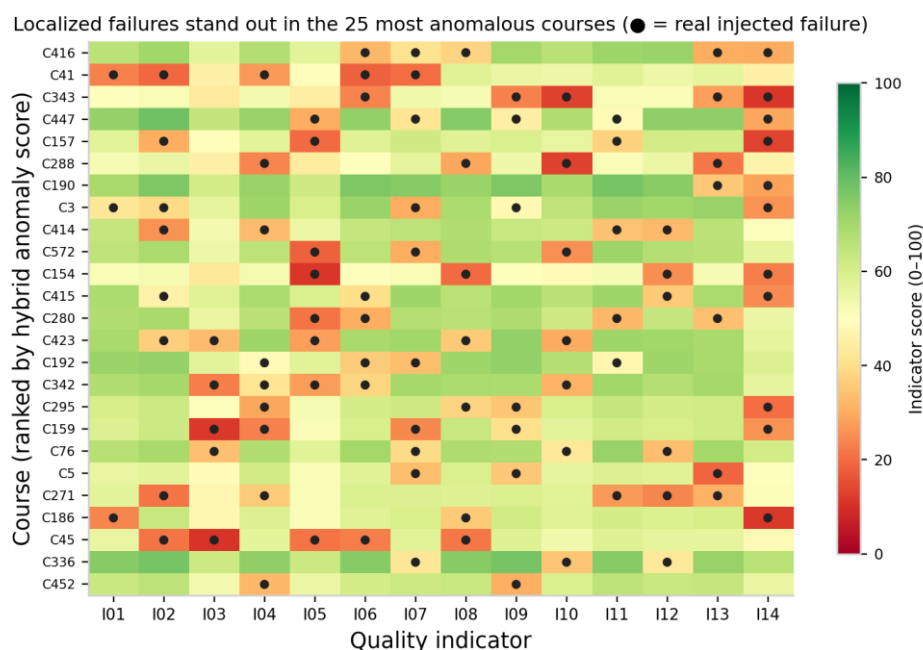


Source: Original figure by the author (2026).



Figure 7

Evaluation sub-matrix of the 25 programmes with the highest hybrid anomaly scores (faithful re-execution of the protocol). How to read: each row is a programme (sorted by anomaly score) and each column is one of the 14 indicators; green cells are healthy scores and red cells are depressed scores. The ● marks identify the real injected failures: the localized red “stains” indicators abruptly depressed in otherwise average programmes are exactly the failure pattern that aggregate means tend to dilute.



Source: Original figure by the author (2026).

DISCUSSION

ENGINEERING AS THE ROBUST STRAND OF QUALITY ASSURANCE

The 4th phase of INAAREES external evaluation, dedicated to STEM programmes mostly engineering and technologies constitutes the largest evaluation block of the national process: 241 programmes from 70 institutions, examined by 445 specialists [17]. The official per-phase aggregates (Table 7, Figure 8) reveal that engineering and STEM programmes show an accreditation rate of 63.5%, above the national average of 62.3% and clearly higher than that of HASS programmes (53.1%), whose non-accreditation rate (46.9%) is the highest of the process [16]. In terms of the proposed matrix model, engineering programmes concentrate in the “robust” and “consolidating” latent profiles, exhibiting lower spectral dispersion a robust strand of the national quality assurance landscape.



Table 7

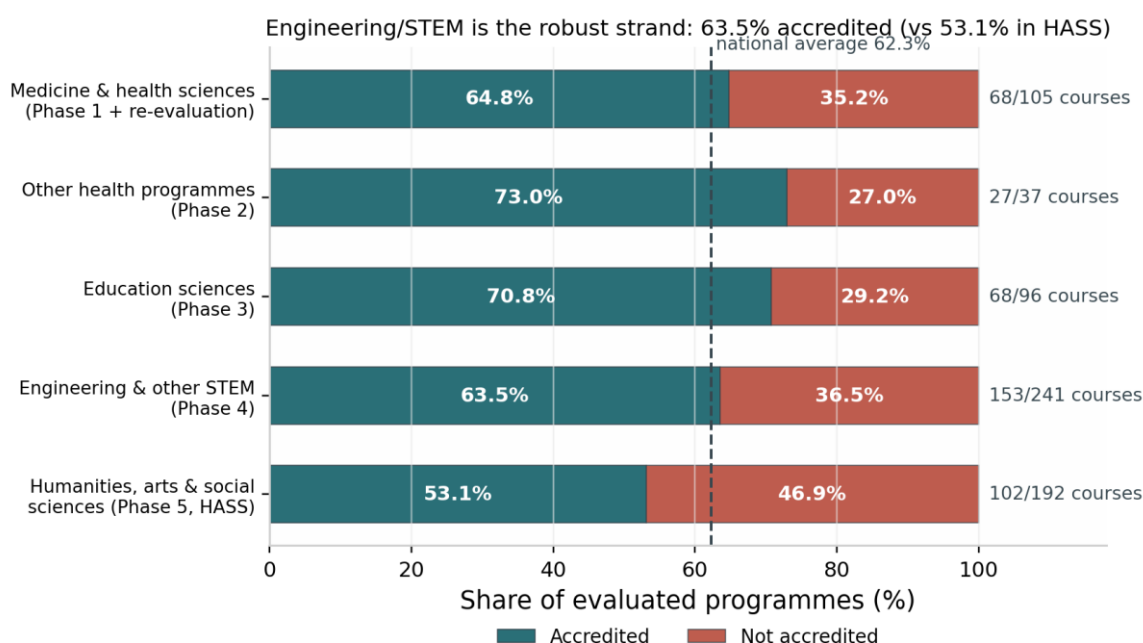
Official aggregate results of INAAREES external evaluation, by phase/area (2023–2026) [16, 17].

Evaluated area	Programmes evaluated	Accredited	Not accredited
Medicine and health sciences (1st phase + re-evaluation)	105	68 (64.8%)	37 (35.2%)
Other health programmes (2nd phase)	37	27 (73.0%)	10 (27.0%)
Education sciences (3rd phase)	96	68 (70.8%)	28 (29.2%)
Engineering and other STEM (4th phase)	241	153 (63.5%)	88 (36.5%)
Humanities, arts and social sciences (5th phase, HASS)	192	102 (53.1%)	90 (46.9%)
Total	671	418 (62.3%)	253 (37.7%)

Source: Original figure by the author (2026).

Figure 8

Official results of INAAREES external evaluation by phase/area (2023–2026) [16, 17]. How to read: each horizontal bar totals 100% of the programmes evaluated in that phase, split into accredited (teal) and not accredited (red); the dashed line marks the national average accreditation rate (62.3%). Bars ending to the right of the line outperform the system — note that engineering/STEM (63.5%) sits above the average while HASS (53.1%) records the lowest accreditation rate of the process. The right-hand labels give the absolute counts.



Source: Original figure by the author from official public aggregates (2026).



This robustness of engineering education is substantive, not accidental. Engineering programmes operate with consolidated traditions of international accreditation and outcome-based education built on measurable learning results [19, 20], which produces more structured and internally coherent evidence matrices. From the standpoint of the present method, this has two practical consequences. First: the low-rank structure of the evaluation matrix is partly sustained by the engineering programmes, whose regular profiles define the dominant subspace; it is precisely against that subspace that the RSDF and DICA-R algorithms measure deviations. Second: the smaller number of anomalies in engineering frees up scrutiny capacity for the areas where risk concentrates — a reading that the IRF makes operational by jointly weighing anomaly and non-compliance. In short, in the model, engineering functions as the system's stability reference axis: without that axis, the spectral residual would no longer have a “regular behaviour” against which to measure failures.

IMPLICATIONS FOR THE INAAREES EXTERNAL EVALUATION PROCESS

The experimental results support three concrete operational applications, compatible with the process regulated by Executive Decree No. 109/20 [15]:

- (a) Prior triage of external evaluation committees.** Before committees are formed and site visits conducted, RSDF+DICA-R can process the self-assessment evidence matrix and flag, in seconds, the programmes whose indicators diverge from the systemic pattern steering evaluators' effort towards where the probability of non-conformity is highest. In the phases carried out, each cycle mobilized hundreds of external evaluators, including international experts [17]; computational triage reduces costs and increases the density of findings per visit.
- (b) Coherence auditing of self-declared evidence.** DICA-R was designed to detect internal inconsistencies between indicator families the typical pattern of inflated data or artificial documentary evidence. In a system where evaluation combines self-declaration, documentary analysis and site visits, an automatic inconsistency alarm is an instrument of process integrity.



(c) Inter-cycle risk management. The IRF provides a risk classification framework that complements the legal accreditation classification provided for in the INAAREES Organic Statute [14]: while the latter reports the outcome of a cycle, the IRF makes it possible to monitor the risk trajectory between cycles, supporting preventive corrective measures a competence expressly assigned to the institute [14].

WHY RANDOMIZE?

At the current dimension of the Angolan process, the advantage of randomization is not speed 671 $\times 14$ matrices are factorized instantly by exact methods (Table 3) but three structural properties: (i) guaranteed scalability for the longitudinal, disaggregated databases the system will inevitably accumulate (the $37.9\times$ speedup at $20,000 \times 2,000$ anticipates that scenario); (ii) statistical stability: averaging randomized repetitions works as a regularization mechanism, barely sensitive to perturbations of a single factorization; and (iii) probabilistic auditability: the RandNLA guarantees (Eq. 4) make it possible to state, with mathematical rigour, the error margin of the processing an institutionally relevant argument for a regulator whose decisions have legal consequences for institutions [12].

LIMITATIONS

Three limitations must be clearly stated. First: the empirical validation was carried out on synthetic data; although calibrated with official aggregates and designed with interpretable profiles, real INAAREES data may exhibit structures not captured by the four-factor model (for example, effects of province, scientific area or the institution's legal regime). Second: the 14 proposed indicators are the author's operationalization, not the official indicator framework of the External Evaluation Manual [9]; transferring to the official framework may require reweighting. Third: the IRF weights (0.6/0.4) were fixed a priori and not optimized; in a real pilot they should be calibrated against historical committee decisions.



ETHICAL CONSIDERATIONS

The proposed system is a decision-support instrument, never an automatic decision-maker: no accreditation, corrective measure or sanction may follow from a computational score without adversarial proceedings, under the terms of the legal evaluation process [12, 15]. The use of institutional data must observe data protection and the confidentiality of ongoing accreditation processes. Finally, algorithmic transparency all algorithms in this article are published in full is a condition of legitimacy: evaluated institutions have the right to know the logic of the instruments that precede them in scrutiny.

CONCLUSION AND FUTURE WORK

This article presented the first study, in Angola, of the application of randomized numerical linear algebra to higher education quality assurance. Starting from the recognition that Africa's educational question is today, emphatically, a crisis of quality rather than quantity, we formalized INAAREES external evaluation as a matrix analysis problem and demonstrated, in a self-authored experimental protocol with controlled ground truth, that: (i) the evaluation matrix has an effectively low-rank structure (99.78% of energy in four factors), which legitimizes randomized methods; (ii) the randomized SVD reproduces the Eckart–Young optimum with a maximum deviation of 0.003 percentage points and scales up to $37.9\times$ faster; (iii) the hybrid RSDF+DICA-R detector fully recovered the 60 injected failures (100% precision at the operational cut-off; AUC-PR = 0.976), including mild failures of one or two indicators; (iv) the IRF prioritizes institutions with a correlation of 0.633 with the true failure rate, additionally capturing structural fragility; and (v) engineering programmes, with 63.5% accreditation in the STEM phase, constitute the robust strand of the system, serving as the reference axis against which the algorithms measure deviations.

As future work, the following stand out: (i) establishing a protocol with INAAREES for a pilot with anonymized real data from one evaluation cycle; (ii) temporal extension of the model (three-mode matrices: programmes $\times$ indicators $\times$ cycles, with randomized tensor decompositions); (iii) integration of



the IRF into the institute's statistical information system; and (iv) replication of the method by other African quality assurance agencies, in the harmonization spirit of the HAQAA initiative [17]. The author makes the algorithms and the experimental protocol available to the scientific community and to Angolan higher education institutions, in particular the Universidade Gregório Semedo, the Universidade Metodista de Angola and the Universidade Lusíada de Angola, as a contribution to a technologically more capable national quality assurance system.

STATEMENT OF ORIGINALITY AND DATA

The algorithms (SRVD, RSDF, DICA-R, IRF) and the experimental method are the author's own work; all charts, tests, tables and computations presented in this article are the author's own elaboration. The empirical component uses a synthetic matrix calibrated with official public INAAREES aggregates; no real institutional data were used. Conflict of interest: the author declares no conflict of interest. Funding: not applicable.

REFERENCES

- [1] Halko, N., Martinsson, P.-G., & Tropp, J. A. (2011). Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2), 217–288.
- [2] Mahoney, M. W. (2011). Randomized algorithms for matrices and data. *Foundations and Trends in Machine Learning*, 3(2), 123–224.
- [3] Woodruff, D. P. (2014). Sketching as a tool for numerical linear algebra. *Foundations and Trends in Theoretical Computer Science*, 10(1–2), 1–157.
- [4] Drineas, P., & Mahoney, M. W. (2016). RandNLA: Randomized numerical linear algebra. *Communications of the ACM*, 59(6), 80–90.
- [5] Martinsson, P.-G., & Tropp, J. A. (2020). Randomized numerical linear algebra: Foundations and algorithms. *Acta Numerica*, 29, 403–572.



- [6] Tropp, J. A., Yurtsever, A., Udell, M., & Cevher, V. (2017). Practical sketching algorithms for low-rank matrix approximation. *SIAM Journal on Matrix Analysis and Applications*, 38(4), 1454–1485.
- [7] Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3), 211–218.
- [8] Johnson, W. B., & Lindenstrauss, J. (1984). Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26, 189–206.
- [9] INAAREES. (2022). *Manual de avaliação externa de cursos e/ou programas* [External evaluation manual for courses and/or programmes] (1st ed.). Damer Gráficas. (In Portuguese)
- [10] Mambo, A. Q. (2025). Ensino superior em Angola: breve análise comparativa dos resultados da avaliação externa [Higher education in Angola: a brief comparative analysis of external evaluation results]. *Sapientiae — Revista Científica da Universidade Oscar Ribas*.
<https://doi.org/10.37293/sapientiae111.10> (In Portuguese)
- [11] Tomé, J. (2024). Garantia da qualidade no ensino superior em Angola: percurso, desafios do presente e do futuro, e o papel das instituições de ensino superior nesse processo [Quality assurance in Angolan higher education: trajectory, present and future challenges, and the role of higher education institutions]. In *XXXII Encontro da Associação das Universidades de Língua Portuguesa, São Tomé and Príncipe*. <https://doi.org/10.31492/978-989-8271-22-8.ATAS2023> (In Portuguese)
- [12] Republic of Angola. (2018). *Presidential Decree No. 203/18 of 30 August* — Legal Regime for the Evaluation and Accreditation of the Quality of Higher Education Institutions. Diário da República, Luanda. (In Portuguese)
- [13] Republic of Angola. (2020). *Presidential Decree No. 310/20 of 7 December* — Legal Regime of the Higher Education Subsystem. Diário da República, Luanda. (In Portuguese)



- [14] Republic of Angola. (2013). *Presidential Decree No. 172/13 of 29 October* — Organic Statute of the National Institute for Assessment, Accreditation and Recognition of Higher Education Studies (INAAREES). *Diário da República*, Luanda. (In Portuguese)
- [15] Republic of Angola. (2020). *Executive Decree No. 109/20 of 10 March* — Regulation of the External Evaluation and Accreditation Process for Courses and/or Programmes and HEIs (amended by Executive Decree No. 148/24 of 6 August). *Diário da República*, Luanda. (In Portuguese)
- [16] INAAREES. (2026). *Ceremony for the release of the results of the external evaluation and accreditation process in higher education: STEM programmes (4th phase), medicine and other health sciences (re-evaluation) and HASS programmes (5th phase)* [Institutional presentation]. Ministry of Higher Education, Science, Technology and Innovation, Luanda.
- [17] Ministry of Higher Education, Science, Technology and Innovation. (2025–2026). *National process of external evaluation and accreditation of higher education quality: STEM and HASS phases* [Institutional press releases]. Luanda.
- [18] UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. UNESCO Publishing.
- [19] International Engineering Alliance. (2021). *Graduate attributes and professional competencies* (4th ed.). IEA Washington Accord, Sydney Accord, Dublin Accord.
- [20] ABET Engineering Accreditation Commission. (2024). *Criteria for accrediting engineering programs, 2024–2025*. ABET, Baltimore, MD.