

RACE-SDG: A ROBUST ASSURANCE-CALIBRATED EXPLAINABLE AI FRAMEWORK FOR SDG-ALIGNED SUSTAINABILITY ACCOUNTING

Jayaprakasha P C

Department of Mathematics, Government First Grade College for Women,
Tumakuru, Karnataka, India.

ORCID: <https://orcid.org/0009-0005-4263-9640>; Email: jayagagandeep2@gmail.com

Abstract

Purpose. In reality, sustainability scores hide three mathematically separate problems: different indicators have both incompatible units and directions; values are estimated or extracted, and so uncertain; a high aggregate can mask the poor performance of a weak but material indicator. This chapter develops RACE-SDG, a general-purpose explainable artificial-intelligence assurance-calibrated framework that produces an auditable SDG-aligned score of the sustainability profile of corporations and does not treat model output as accounting fact.

Design/methodology/approach. Data as of October 2023 are employed for direction-aware logarithmic gradual progression mapped to a bounded scale; materiality and empirical data pooled using entropy-based weights; machine-learning predictions obtained while using split-conformal intervals. An ellipsoid represents uncertainty in the indicator. Due to convex duality the resulting worst-case weighted score has closed form. A generalized power mean gives a non-compensatory diagnostic and indicator-level uncertainty allocations reconcile exactly to the robust score. The numerical application draws upon independently assured FY2023-FY2024 Business Responsibility and Sustainability Reporting data published by Infosys Limited; shock parameters are designated independently and are not shown as empirical data.

Findings. The unadjusted nominal score for the axial scores on the direction is 0.6976 for seven indicators. This is reduced to 0.6597 with a one-sided 95% uncertainty radius, and the weighted geometric mean is 0.6638. RACE-SDG combined score: 0.6613 The result is most positively affected by emissions intensity, renewable-energy share and waste recovery, and negatively by energy intensity and the female wage share. For SDG 12, the SDG-level scores are 0.6068 (SDG7), 0.5085 (SDG8), 0.7025(SDG12), and 0.6893(SDG13).

Originality/value. This includes synthesizing predictive AI with finite-sample calibration, robust optimization, non-compensatory aggregation, and a reconciliation of uncertainty like one would find in an accounting. Each aggregate can be assigned to disclosed values, a direction convention, a weight of materiality and an uncertainty charge explicitly.

Keywords: sustainability accounting; explainable artificial intelligence; robust optimization; SDG reporting; conformal ESG scoring; prediction; BRSR.

1. Introduction

Sustainability accounting is no longer merely a descriptive appendix to financial reporting. It now enables capital allocation, operational governance, risk management and communication with regulators and stakeholders. This expanded role has crystallised an age-old differentiator; disclosure is not measurement and measurement is certainly not associated with decision usefulness. Sustainable Development — Sustainability accounting must not only be relevant for the accountability of an

organisation, but it also has to provide information relevant for managerial decision making (Burritt and Schaltegger, 2010); Maas et al. (2016) find that assessment, management control and external reporting reveal weak integration. Not unlike reviews of accounting practice, some identify consistent variation in the adoption, scope and quality of reporting (Hahn & Kühnen 2013; Christensen et al., 2021).

Data is transformed into sensors, enterprise records, supplier files and narrative reports using artificial intelligence on a scale that no one can imagine by manual reporting. It is capable of predicting missing values, future paths, outlier detection (surfacing unusual records), and even rationalization of nonlinear relationships. AI is now seen as an important tool for achieving the Sustainable Development Goals (SDGs), but these effects are not all positive (Vinuesa et al. Machine learning can help with mitigation in every sector — energy, transport, buildings, industry, and finance (Rolnick et al., 2022; Kaack et al., 2022) but the computational system itself is resource-demanding. Efficiency and transparent reporting of computational cost (Schwartz et al., 2020; Strubell et al., 2019) is thus a requirement for GreenAI.

False precision is the core risk on accounting. As the agencies vary in scope, measurement and aggregation of ESG rating (Berg et al., 2022), therefore there is divergence along this index. Experiencing problems related to incomplete histories, evolving definitions and changing boundaries due to different choices of estimation (Kotsantonis & Serafeim 2019), are among the common pitfalls encountered by data vendors also. For instance, the methodologies may underestimate good sustainability principles or allow for trade-offs across the environmental and social dimensions (Escrig-Olmedo et al., 2019). These constraints are significant, as investors apply ESG facts for valuation and governance (Amel-Zadeh & Serafeim, 2018), cloth sustainability troubles have specific performance implications (Khan et al., 2016), and shape can affect market perceptions concerning underlying performance (Mervelskemper & Streit, 2017). Evidence of a connection between ESG and financial performance has been the subject of extensive investigation: Individual studies are abundant yet inconsistent concerning the results (Friede et al., 2015; Eccles et al., 2014; Fatemi et al.

This chapter gives contracted answers to; how is it possible for AI assisted sustainability accounting to create a score which can be compared, is uncertainty-aware, future-proof against compensation and fully reconcilable? It accomplishes this through five interrelated components RACE-SDG. The problem is two-fold, firstly a log transform maintains the direction of each indicator and supports a proportional scale. Weights depend on the relevance of a topic, the kind of information and quality of assurance. Thirdly, we can derive distribution-free conformal bounds on forecasts made by machine learning models. Fourth, the upper bound on a worst-case score can be mapped from those bounds back into an ellipsoidal uncertainty set. Fifth, we demonstrate imbalance and uncertainty in a power-mean diagnostic with an additive contribution ledger.

It is a contribution of mathematics, not rhetoric. This score being proposed, is not a new proprietary rating with an ambiguous meaning. A hypervisible assumption, an inferred strong counterpart, a numeric output that could be mechanically recomputed from the disclosed inputs. This design matches the framework propounded in explainable-AI work emphasizing local fidelity (Ribeiro et al., 2016), global interpretability (Guidotti et al., 2018; Barredo Arrieta et al., 2020) and governance proportional to consequential decisions (Lundberg et al., 2020).

2. Research Background and Gap

2.1 Measurement boundaries and comparability

Environmental indicators are boundary-dependent. A carbon footprint differs depending on whether or not you include purchased electricity, business travel, suppliers and embodied emissions. Matthews et al. Importance of complete estimation boundaries (2008) and Hertwich & Peters (2009), Peters et al. (2011) demonstrate how trade reallocates production and consumption-based emissions. Eora and other multi-region input-output systems support wider value-chain estimation (Lenzen et al., 2013). An overall score that doesn't take into account what happens to the boundaries could look at reclassification as an improvement.

The model therefore requires a boundary-consistency flag before any arithmetic is performed. Let $b_{ijt} = 1$ when indicator j for entity i at time t is comparable with its reference period, and let $b_{ijt} = 0$ otherwise. A non-comparable observation is not silently imputed into a reported score. It is forecast, bounded, and labelled as estimated. This rule makes AI a controlled estimation layer rather than a substitute for evidence.

2.2 Prediction, anomaly control, and explainability

Tree ensembles are appropriate for analyzing mixed sustainability data as they can deal with nonlinearities, interactions and missingness. Random forests avoid variance via decorrelated trees (Breiman 2001), gradient boosting with functionals takes the robust duty of functional gradient descent (Friedman 2001) and XGBoost adds regularization and scalable computation (Chen et Guestrin 2016). For unsupervised anomaly detection, Isolation Forest is able to examine unusual records without a required labelled anomaly class (Liu et al., 2008), while support-vector data description offers an alternative one-class boundary (Tax & Duin, 2004). These techniques can be used to identify records for audit, however an anomalous score does not mean an error, it simply signals a risk warranting documentary review.

Explanations are equally important. An example of this might be the work on LIME, which builds a local surrogate around a prediction (Ribeiro et al., 2016); tree-based Shapley explanations yield consistent feature attributions (Lundberg et al., 2020); and broader surveys distinguish local/global, model-specific/model-agnostic and intrinsic/post-hoc explanation (Guidotti et al., 2018; Barredo Arrieta et al., 2020). Layer 2: Decomposing the Robust Accounting Score RACE-SDG performs additional explanation layers post-prediction, where it decomposes the final robust accounting score into indicator value and uncertainty charges that exactly add up to the total.

2.3 The unresolved aggregation problem

The dashboards simply normalize the indicators and arrive at the weighted arithmetic mean. An approach that is clear but allows for full compensation — great waste recovery can offset a decreasing energy intensity. An endogenous efficiency frontier can be obtained with data envelopment analysis (Charnes et al., 1978), e.g. slacks-based measures that simultaneously account for input excesses and output shortfalls (Tone, 2001). Fuzzy-set methods can model graded membership when evidence has a linguistic, or imprecise nature (Zadeh 1965). Multiobjective evolution methods can also approximate the trade-off frontiers (Deb et al., 2002). These are great for benchmarking or planning purpose, but score rating that is used in corporate reporting needs fixed set of materiality weights and auditable reconciliation.

RACE-SDG retains the interpretation of weighted means, integrates a principled uncertainty deduction and offers a non-compensatory diagnostic for robust generalization power: a generalized power mean

weights performance on all classifications equally. The two answers: the robust mean asks what score you could still obtain if your measurement is rigged; the power mean asks whether strong dimension is due to hidden weak dimensions.

3. Mathematical Framework

3.1 Indicator direction and proportional progress

Let $x_{ijt} > 0$ denote the reported value of indicator $j = 1, \dots, m$ for entity i at time t . Define $d_j = 1$ for a benefit indicator, where a larger value is desirable, and $d_j = -1$ for a burden indicator, where a smaller value is desirable. Relative progress from $t - 1$ to t is

$$r_{ijt} = d_j \ln \left(\frac{x_{ijt}}{x_{ij,t-1}} \right). \quad (1)$$

The logarithm makes equal proportional changes symmetric. A 10% rise followed by its exact reciprocal fall sums to zero in log space. To obtain a bounded score, define

$$z_{ijt} = \sigma(\kappa_j r_{ijt}) = \frac{1}{1 + \exp(-\kappa_j r_{ijt})}, \quad 0 < z_{ijt} < 1, \quad (2)$$

where $\kappa_j > 0$ controls sensitivity. The neutral score is 0.5, and a z more than 0.5 means improved, while z less than 0.5 means worsen. Equation (2) also does not change when an irrelevant peer enters or leaves the comparison set, unlike min-max normalization.

Proposition 1. For $x_{ijt} > 0$, the mapping in Equation (2) is strictly increasing in a benefit indicator and strictly decreasing in a burden indicator. It satisfies $z(x, x) = 0.5$ and $z(x_t, x_{t-1}) + z(x_{t-1}, x_t) = 1$ when direction is held fixed.

Proof. Since $\sigma'(u) = \sigma(u)[1 - \sigma(u)] > 0$ and $\partial r / \partial x_t = d_j / x_t$, the derivative has sign d_j . Neutrality follows from $\ln(1) = 0$. Reciprocal reversal changes r to $-r$, and $\sigma(u) + \sigma(-u) = 1$.

3.2 Materiality, information, and assurance weights

Let $\pi_j > 0$ be the approved double-materiality prior with $\sum_j \pi_j = 1$. When a peer-period panel is available, empirical discrimination can be measured by Shannon entropy (Shannon, 1948). For normalized positive observations y_{ij} ,

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^n y_{ij}}, \quad e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij}, \quad q_j = \frac{1 - e_j}{\sum_{\ell=1}^m (1 - e_\ell)}. \quad (3)$$

Let $a_j \in (0, 1]$ denote assurance quality: higher values correspond to stronger evidence, stable boundaries, and effective controls. The pooled weight is

$$w_j(\eta) = \frac{\pi_j^{1-\eta} q_j^\eta a_j}{\sum_{\ell=1}^m \pi_\ell^{1-\eta} q_\ell^\eta a_\ell}, \quad 0 \leq \eta \leq 1. \quad (4)$$

The geometric pool avoids letting an indicator enter and dominate only because its entropy is low, despite being highly variable but immaterial. In the absence of a peer panel, we can set $\eta = 0$ and reveal the materiality vector π . The most interesting feature of the split score is that assurance indirectly influences it through w_j and the uncertainty model; you should not afford a cosmetic precision premium for low confidence evidence.

3.3 AI forecast and conformal calibration

For an indicator requiring estimation, let $\mathbf{h}_{ijt}$ contain lagged values, operational drivers, boundary flags, and text-derived features. A gradient-boosted predictor takes the form

$$\hat{x}_{ij,t+1} = f_j(\mathbf{h}_{ijt}) = \sum_{k=1}^K v f_{jk}(\mathbf{h}_{ijt}), \quad (5)$$

with learning rate v . XGBoost estimates the trees by minimizing

$$\mathcal{L}_j = \sum_{s=1}^N \ell(x_{js}, \hat{x}_{js}) + \sum_{k=1}^K \left(\gamma T_k + \frac{\lambda}{2} \|\boldsymbol{\omega}_k\|_2^2 \right), \quad (6)$$

where T_k is the leaf count and $\boldsymbol{\omega}_k$ contains leaf values. A point forecast is not enough for accounting. Using an exchangeable calibration set of size n_c , compute residuals $u_s = |x_s - \hat{x}_s|$ and

$$Q_{1-\alpha} = u_{(k)}, \quad k = \text{ceil}((n_c + 1)(1 - \alpha)). \quad (7)$$

The split-conformal interval is

$$C_{1-\alpha}(\mathbf{h}) = [\hat{x}(\mathbf{h}) - Q_{1-\alpha}, \hat{x}(\mathbf{h}) + Q_{1-\alpha}]. \quad (8)$$

Under exchangeability, for a finite number of samples the marginal coverage is at least $1 - \alpha$ (up to the usual discretization; Lei et al. (2018)). It distinguishes predictive uncertainty from how well the model seems to fit, meaning that an average of the past without uncertainty is not reported as the only possibility.

3.4 Ellipsoidal robust score

Let $\hat{\mathbf{z}} \in (0,1)^m$ contain observed or predicted indicator scores. Correlated uncertainty is represented by

$$\mathcal{U}_\rho = \{\mathbf{z} = \hat{\mathbf{z}} + \mathbf{L}\mathbf{u}: \|\mathbf{u}\|_2 \leq \rho\}, \quad \boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top. \quad (9)$$

The accounting question is conservative: what is the lowest weighted score consistent with the approved uncertainty set?

$$S_{\text{rob}}(\rho) = \min_{\mathbf{z} \in \mathcal{U}_\rho} \mathbf{w}^\top \mathbf{z}. \quad (10)$$

Robust optimization protects a decision against adverse parameter realizations (Ben-Tal & Nemirovski, 2000; Bertsimas & Sim, 2004). Moment-based and Wasserstein distributionally robust models extend this protection to ambiguity in probability laws (Delage & Ye, 2010; Mohajerin Esfahani & Kuhn, 2018). For the reporting problem, Equation (10) has a particularly useful closed form.

Theorem 1. If $\boldsymbol{\Sigma} \succcurlyeq 0$ and $\mathbf{w} \geq 0$, $\mathbf{1}^\top \mathbf{w} = 1$, then

$$S_{\text{rob}}(\rho) = \mathbf{w}^\top \hat{\mathbf{z}} - \rho \sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}. \quad (11)$$

Proof. Substituting Equation (9) into Equation (10) gives $\mathbf{w}^\top \hat{\mathbf{z}} + \min_{\|\mathbf{u}\|_2 \leq \rho} (\mathbf{L}^\top \mathbf{w})^\top \mathbf{u}$. By Cauchy-Schwarz, the second term is bounded below by $-\rho \|\mathbf{L}^\top \mathbf{w}\|_2$. Equality is attained at $\mathbf{u}^* = -\rho \mathbf{L}^\top \mathbf{w} / \|\mathbf{L}^\top \mathbf{w}\|_2$ when the denominator is nonzero. Finally, $\|\mathbf{L}^\top \mathbf{w}\|_2^2 = \mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}$. $\square$

Equation (11) is analogous to a risk-adjusted performance measure. Conditional value-at-risk can replace the ellipsoid when scenario losses are available (Rockafellar & Uryasev, 2000), but the ellipsoidal form is parsimonious and exactly reconcilable.

3.5 Non-compensation, SDG mapping, and reconciliation

For $r \neq 0$, define the weighted power mean

$$M_r(\mathbf{z}) = \left(\sum_{j=1}^m w_j z_j^r \right)^{1/r}, \quad M_0(\mathbf{z}) = \prod_{j=1}^m z_j^{w_j}. \quad (12)$$

For $r < 1$, $M_r \leq M_1$ and a weak indicator has greater influence than under the arithmetic mean. The chapter reports M_0 , the weighted geometric mean. The final score is

$$S_{\text{RACE}} = \beta S_{\text{rob}} + (1 - \beta) M_0, \quad 0 \leq \beta \leq 1. \quad (13)$$

Let $A_{kj} \geq 0$ map indicator j to SDG k . An SDG-specific score is

$$G_k = \frac{\sum_{j=1}^m A_{kj} w_j z_j}{\sum_{j=1}^m A_{kj} w_j}. \quad (14)$$

Finally, define nominal contribution $c_j = w_j \hat{z}_j$ and uncertainty charge

$$u_j = \rho \frac{w_j (\Sigma \mathbf{w})_j}{\sqrt{\mathbf{w}^T \Sigma \mathbf{w}}}. \quad (15)$$

Because $\sum_j u_j = \rho \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}$, the robust score reconciles exactly:

$$S_{\text{rob}} = \sum_{j=1}^m (c_j - u_j). \quad (16)$$

This is the explainability layer needed for assurance: management can see whether a low robust score arises from weak performance, unreliable measurement, or both.

4. Data and Numerical Method

4.1 Source and indicator construction

The application relies upon the Infosys Limited Business Responsibility and Sustainability Report, FY2023-FY2024 (Infosys Limited 2024). According to the report, Deloitte Haskins & Sells LLP carried out an independent audit of the BRSR Core indicators we have used here. Four indicators were chosen because both the periods of the values and their calculation bases were revealed for these four metrics as compared with others. Burden indicators include Scope 1+2, Scope 3, energy and water while benefit indicators include renewable share, waste recovery and female wage share. Renewable share and recovery of waste were computed from the disclosed components.

Table 1. Disclosed values and derived ratios

Indicator	Direction	FY2023	FY2024	Proportional change
Scope 1+2 emissions intensity per rupee	Burden	4.80×10^{-8}	4.10×10^{-8}	-14.583%
Scope 3 emissions intensity per rupee	Burden	1.25×10^{-7}	1.17×10^{-7}	-6.400%
Energy intensity per rupee	Burden	5.12×10^{-7}	5.46×10^{-7}	+6.641%
Renewable-energy share	Benefit	0.47890	0.57866	+20.832%
Water intensity per rupee	Burden	1.55×10^{-6}	1.47×10^{-6}	-5.161%
Waste recovery rate	Benefit	0.48246	0.94625	+96.129%

Indicator	Direction	FY2023	FY2024	Proportional change
Female wages as share of total wages	Benefit	0.293	0.294	+0.341%

Renewable and total energy for FY2023 and FY2024 were 359644 and 750,986 GJ (for reproducibility) and 485753 & 839448 GJ respectively. In tones, the quantities recovered and total waste for FY2023 were 10,160.56 and 21,059.77 respectively, while those in FY2024 were 46,918.51 and 49,583.58 respectively. The information presented in Table 1 consists of ratios calculated directly from the described components.

The waste series requires caution. Total waste for FY2024 was significantly higher for that component — construction and demolition — as compared to FY2023. Thus, although the recovery rate is correct mathematically, its comparability among different threats is not as strong as w.r.t intensity indicators. This is captured in RACE-SDG as a wider uncertainty range rather than dropping the observation.

4.2 Parameters and reproducibility

The sensitivity in Equation (2) is $\kappa_j = 10$ for all indicators. The disclosed materiality vector is

$$\mathbf{w} = (0.20, 0.16, 0.14, 0.14, 0.12, 0.12, 0.12)^T. \quad (17)$$

Because this is a single-company, two-period application, entropy weights are not estimated; Equation (4) is used with $\eta = 0$. The vector in Equation (17) is a transparent analytical parameter, not a claim about Infosys management's own weights.

For the uncertainty stress test, $\Sigma = \mathbf{B}\mathbf{B}^T + \text{diag}(\mathbf{d}^2)$, with three factor columns representing climate, resource, and social dependence:

$$\mathbf{B} = \begin{bmatrix} .025 & 0 & 0 \\ .030 & .010 & 0 \\ .025 & .015 & 0 \\ .020 & .010 & 0 \\ 0 & .025 & 0 \\ 0 & .040 & 0 \\ 0 & 0 & .020 \end{bmatrix}, \quad \mathbf{d} = (.020, .025, .020, .020, .025, .050, .020)^T. \quad (18)$$

These are governance-set score-space stress scales. They are not statistical standard errors estimated from two observations. The larger waste idiosyncratic scale records the boundary/composition concern noted above. The primary robust calculation uses $\rho = 1.645$, the conventional one-sided 95% Gaussian quantile, and Equation (13) uses $\beta = 0.60$.

5. Results and Discussion

5.1 Direction-adjusted scores

Substitution into Equations (1)-(2) gives the results in Table 2. For example, Scope 1+2 intensity yields

$$r_1 = -\ln\left(\frac{4.10 \times 10^{-8}}{4.80 \times 10^{-8}}\right) = 0.157629, \quad z_1 = \frac{1}{1 + e^{-10(0.157629)}} = 0.828678. \quad (19)$$

Renewable energy gives $r_4 = \ln(0.57866/0.47890) = 0.189228$ and $z_4 = 0.869015$. Energy intensity moved in the adverse direction, producing $r_3 = -\ln(5.46/5.12) = -0.064294$ and $z_3 = 0.344581$.

Table 2. Transformed scores and robust reconciliation

Indicator	r_j	z_j	$w_j z_j$	u_j at $\rho = 1.645$	$w_j z_j - u_j$
Scope 1+2 intensity	0.157629	0.828678	0.165736	0.006897	0.158839
Scope 3 intensity	0.066140	0.659574	0.105532	0.008141	0.097390
Energy intensity	-0.064294	0.344581	0.048241	0.006523	0.041718
Renewable-energy share	0.189228	0.869015	0.121662	0.005073	0.116589
Water intensity	0.052993	0.629466	0.075536	0.003409	0.072127
Waste recovery rate	0.673604	0.998814	0.119858	0.006998	0.112860
Female wage share	0.003407	0.508517	0.061022	0.000823	0.060199
Total			0.697587	0.037864	0.659722

An unchanged metric does not tick the box for failure in the score: no change maps to 0.5. It instead measures direction-adjusted momentum. Thus, this application should be interpreted as a progress score—not a static sustainability-level score. One way to include absolute target attainment is by substituting a science-based or regulatory target for the denominator in Equation (1).

5.2 Robustness and non-compensation

From Equations (17)-(18),

$$\sqrt{\mathbf{w}^T \Sigma \mathbf{w}} = 0.023018. \quad (20)$$

The nominal score is $S_{\text{nom}} = 0.697587$. At $\rho = 1.645$, the uncertainty charge is $1.645(0.023018) = 0.037864$, so $S_{\text{rob}} = 0.659722$. Table 3 shows that the conclusion is not dependent on a single radius.

Table 3. Robust-score sensitivity

Radius ρ	Interpretation	$S_{\text{rob}}(\rho)$
0.000	Nominal, no uncertainty deduction	0.697587
0.500	Mild stress	0.686078
1.000	Unit-radius stress	0.674569
1.645	One-sided 95% reference radius	0.659722
2.000	Severe stress	0.651551

The weighted geometric mean is

$$M_0 = \prod_{j=1}^7 z_j^{w_j} = 0.663767, \quad (21)$$

below the arithmetic mean because the adverse energy score and near-neutral wage-share score are not fully offset by waste recovery. The combined result is

$$S_{\text{RACE}} = 0.60(0.659722) + 0.40(0.663767) = 0.661340. \quad (22)$$

This specific value is very precise: 66.13% of the bounded progress scale remains after an approved, correlated uncertainty stress from a nn-transformation step plus a non-compensation check. A separately validated (decision rule) should not translate that into an investment rating.

5.3 SDG alignment

The crosswalk assigns: energy intensity and renewable share to SDG 7; female wage share to SDG 8; energy, renewable energy, water and waste to SDG 12; Scope 1+2, Scope 3, energy and renewable energy to SDG 13. Equation (14) yields Table 4.

Table 4. SDG-level progress scores

SDG	Included accounting signals	G_k
SDG 7 - Affordable and Clean Energy	Energy intensity; renewable-energy share	0.606798
SDG 8 - Decent Work and Economic Growth	Female wage share	0.508517
SDG 12 - Responsible Consumption and Production	Energy; renewable energy; water; waste	0.702494
SDG 13 - Climate Action	Scope 1+2; Scope 3; energy; renewable energy	0.689330

This also explains why separate indicators are needed for SDG 7. The composition of the activity, renewables share improved well and energy intensity deteriorated. Thus, any narrative that only claims renewable utilization accounted for a higher portion would be hiding an important contra signal. Since the reported percentage of female gross wages only increased from 29.3% to 29.4%, SDG 8 score neutral where such minor variation does not generate a strong case for social performance which RACE-SDG would otherwise amplify.

5.4 Assurance interpretation

The highest uncertainty fees are Scope 3 depth (1, and waste recuperation, scope 1+2 and just underneath them power depth (0.006523). These costs include autonomous unpredictability and covariance with other weighted signs. The achievement is counterbalanced: against net dues (0.697587) the expenses (0.037864) balance out an extraordinary score of 0.659722.

It's a mechanism that is an improvement over a hidden haircut. It tells an assurer which observations make scores fragile. Equation (11), which integrates improved covariance matrix with evidence, reduces the charge if tighter Scope 3 uncertainty exists due to better supplier evidence. If the underlying performance doesn't change, the model is simply better evidence that assurance actually does more to mitigate emissions without preventing their power.

6. Implementation Protocol

Provably via governance isolation: Guaranteeing an immutable evidence-to-score trail in production by design. Algorithms can and do provide relevant sensitivity assessments and optimal trade-offs but we cannot solely rely on algorithms to determine social or environmental materiality without some form of responsible human guardianship. In settings of a fixed transition budget, the score can become an objective or constraint in a Mult objective program. Cost-emissions-social trade-offs are investigated using appropriate Pareto methods including NSGA-II (Deb et al., 2002), and the robust counterparts that enable feasibility to be preserved when abatement coefficients remain uncertain (Bertsimas & Sim, 2004).

7. Limitations and Research Directions

The application has four limitations. In the first step, only two periods are insufficient to learn the AI component or compute empirical covariance. Specifically, the numerical section portrays scoring and stress-test layers with pre-disclosed data; whereas the forecasting equations paragraph specifies various extensions to production. Also, mound/limits on the weights and levels of uncertainty are observable analytical parameters, not company-indorsed disclosures about what is material. To finish, the score measures change over time from the end of last reporting period to now. Over distance from a baseline head above water but not swimmable means target-based distance must move with momentum as operationally used. Fourth, results from a single firm do not provide prescriptions for performance.

Future work would: estimate multiple-firm BRSR panel data, check stability under restatements, and contrast conformal coverage by sectors. Distributionally robust formulations (Delage & Ye 2010; Mohajerin Esfahani & Kuhn 2018) could be previously substituted for some ellipsoids with nonstationary tails of supplier data. However, the efficiency of slacks-based peer benchmarks may be leveraged in our analysis with restrictions on the endogeneity of dispersion weights (Tone, 2001). Within this context, calibrated fuzzy membership functions could represent linguistic evidence in situations where the quantitative assessment is unequivocally inadequate (Zadeh 1965). What constitutes the central empirical test here is whether conventional arithmetic of uncertainty-reconciled scores is more reliable at predicting restatements, assurance findings, delivery against transition-plans or financing outcomes.

8. Conclusion

Sustainable accounting and materiality, AI cannot circumvent prediction / uncertainty; disentangling those mathematically is key before it adds value (Letter from sceptic) RACE-SDG provides that separation. It employs proportional improvement around bounded scores, materiality and evidence quality pooling via Bansal & Hwang (2021), forecast calibration – literally subtracts a closed-form-correlated uncertainty charge, uses a power mean for compensation (low variance compensates high variance aggregate), maps output onto SDGs and reconciles summation with indicator contributions.

Application for FY2023-FY2024 generates a nominal score of 0.6976, a robust score of 0.6597 and final RACE-SDG score of 0.6613. The exact number is meaningless compared to the audit trail behind it. The positive signals of renewable-energy and waste-recovery run together with negative energy intensity and an effective standstill of female wage share. Instead of squashing the tensions into some $x y$ coordinate that we cannot yet articulate, the framework contains those tensions. What this characteristic certainly makes it perfectly suited for, however, is in the area of AI-enabled corporate reporting in which credibility depends not only on computation but also on how evidence, judgment and uncertainty made their way into the number we report.

References

- [1.] Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, 74(3), 87-103. <https://doi.org/10.2469/faj.v74.n3.2>
- [2.] Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bannetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- [3.] Ben-Tal, A., & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88, 411-424. <https://doi.org/10.1007/PL00011380>

- [4.] Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315-1344. <https://doi.org/10.1093/rof/rfac033>
- [5.] Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35-53. <https://doi.org/10.1287/opre.1030.0065>
- [6.] Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [7.] Burritt, R. L., & Schaltegger, S. (2010). Sustainability accounting and reporting: Fad or trend? *Accounting, Auditing & Accountability Journal*, 23(7), 829-846. <https://doi.org/10.1108/09513571011080144>
- [8.] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- [9.] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>
- [10.] Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26, 1176-1248. <https://doi.org/10.1007/s11142-021-09609-5>
- [11.] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197. <https://doi.org/10.1109/4235.996017>
- [12.] Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3), 595-612. <https://doi.org/10.1287/opre.1090.0741>
- [13.] Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835-2857. <https://doi.org/10.1287/mnsc.2014.1984>
- [14.] Escrig-Olmedo, E., Fernández-Izquierdo, M. Á., Ferrero-Ferrero, I., Rivera-Lirio, J. M., & Muñoz-Torres, M. J. (2019). Rating the raters: Evaluating how ESG rating agencies integrate sustainability principles. *Sustainability*, 11(3), 915. <https://doi.org/10.3390/su11030915>
- [15.] Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal*, 38, 45-64. <https://doi.org/10.1016/j.gfj.2017.03.001>
- [16.] Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233. <https://doi.org/10.1080/20430795.2015.1118917>
- [17.] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [18.] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), Article 93. <https://doi.org/10.1145/3236009>
- [19.] Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59, 5-21. <https://doi.org/10.1016/j.jclepro.2013.07.005>
- [20.] Hertwich, E. G., & Peters, G. P. (2009). Carbon footprint of nations: A global, trade-linked analysis. *Environmental Science & Technology*, 43(16), 6414-6420. <https://doi.org/10.1021/es803496a>
- [21.] Infosys Limited. (2024). *Business Responsibility and Sustainability Report 2023-24*. <https://www.infosys.com/investors/reports-filings/annual-report/annual/documents/2023-24/business-responsibility-and-sustainability-report.pdf>
- [22.] Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12, 518-527. <https://doi.org/10.1038/s41558-022-01377-7>
- [23.] Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91(6), 1697-1724. <https://doi.org/10.2308/accr-51383>

- [24.] Kotsantonis, S., & Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, 31(2), 50-58. <https://doi.org/10.1111/jacf.12346>
- [25.] Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R. J., & Wasserman, L. (2018). Distribution-free predictive inference for regression. *Journal of the American Statistical Association*, 113(523), 1094-1111. <https://doi.org/10.1080/01621459.2017.1307116>
- [26.] Lenzen, M., Moran, D., Kanemoto, K., & Geschke, A. (2013). Building Eora: A global multi-region input-output database at high country and sector resolution. *Economic Systems Research*, 25(1), 20-49. <https://doi.org/10.1080/09535314.2013.769938>
- [27.] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413-422). <https://doi.org/10.1109/ICDM.2008.17>
- [28.] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- [29.] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2, 56-67. <https://doi.org/10.1038/s42256-019-0138-9>
- [30.] Maas, K., Schaltegger, S., & Crutzen, N. (2016). Integrating corporate sustainability assessment, management accounting, control, and reporting. *Journal of Cleaner Production*, 136, 237-248. <https://doi.org/10.1016/j.jclepro.2016.05.008>
- [31.] Matthews, H. S., Hendrickson, C. T., & Weber, C. L. (2008). The importance of carbon footprint estimation boundaries. *Environmental Science & Technology*, 42(16), 5839-5842. <https://doi.org/10.1021/es703112w>
- [32.] Mervelskemper, L., & Streit, D. (2017). Enhancing market valuation of ESG performance: Is integrated reporting keeping its promise? *Business Strategy and the Environment*, 26(4), 536-549. <https://doi.org/10.1002/bse.1935>
- [33.] Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171, 115-166. <https://doi.org/10.1007/s10107-017-1172-1>
- [34.] Peters, G. P., Minx, J. C., Weber, C. L., & Edenhofer, O. (2011). Growth in emission transfers via international trade from 1990 to 2008. *Proceedings of the National Academy of Sciences*, 108(21), 8903-8908. <https://doi.org/10.1073/pnas.1006388108>
- [35.] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144). <https://doi.org/10.1145/2939672.2939778>
- [36.] Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21-41. <https://doi.org/10.21314/JOR.2000.038>
- [37.] Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., & Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, 55(2), Article 42. <https://doi.org/10.1145/3485128>
- [38.] Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54-63. <https://doi.org/10.1145/3381831>
- [39.] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [40.] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3645-3650). <https://doi.org/10.18653/v1/P19-1355>
- [41.] Tax, D. M. J., & Duin, R. P. W. (2004). Support vector data description. *Machine Learning*, 54, 45-66. <https://doi.org/10.1023/B:MACH.0000008084.60811.49>

- [42.] Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3), 498-509. [https://doi.org/10.1016/S0377-2217\(99\)00407-5](https://doi.org/10.1016/S0377-2217(99)00407-5)
- [43.] Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11, 233. <https://doi.org/10.1038/s41467-019-14108-y>
- [44.] Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)