Engineering and Technology Journal e-ISSN: 2456-3358

Volume 11 Issue 08 August-2026, Page No.-11123-11132

DOI: 10.47191/etj/v11i08.06, I.F. – 8.595

Bridging Standards and Sensors: A Systematic Literature Review and Quantitative Evidence Synthesis of Carbon Accounting Methods for Energy Efficiency Across Terrestrial, Marine, And Built-Environment Ecosystems

Eka Sudarmaji

Fakultas Ekonomi dan Bisnis, Universitas Pancasila

ABSTRACT

Background: Carbon accounting has moved from a specialist scientific discipline into a mainstream requirement of corporate governance, financial regulation, and ecosystem management. The literature measuring it, however, remains divided between two largely non-communicating traditions: administrative and disclosure-based accounting on one hand, and physically measured, sensor-based accounting on the other.

Objective: This study systematically reviews the governance and standards literature on carbon accounting and synthesises reported measurement-accuracy metrics from primary studies spanning forest, soil, marine, agricultural, built-environment, and corporate domains, with the aim of characterising the degree of alignment — or misalignment — between these two traditions and their implications for energy-efficiency-linked policy.

Methods: A systematic literature review with quantitative evidence synthesis was conducted following the PRISMA 2020 framework. Searches in Scopus and Web of Science used the Boolean string, restricted to English-language peer-reviewed sources from 2011 to June 2026. After deduplication and screening, 79 unique records were retained: 35 governance and standards sources synthesised narratively, and 55 quantitative primary studies (60 statistical records) pooled descriptively by ecosystem and sector domain. Formal inverse-variance-weighted meta-analysis was not performed due to outcome-measure heterogeneity; accuracy statistics (R2, RMSE, rRMSE) are reported as domain-level ranges. A five-criterion quality appraisal was applied to all quantitative records, and a sensitivity analysis examined the effect of excluding China-based studies on domain-level accuracy estimates.

Results: Model accuracy varied sharply by domain. Engineered and semi-controlled systems — UAV-LiDAR forest biomass inversion (R2 = 0.93-0.95) and residential building carbon-intensity modelling (R2 = 0.91) — consistently outperformed open-ecosystem reconciliation models, where regional forest-carbon inversions fell as low as R2 = 0.147. Sensitivity analysis showed that excluding China-based forest studies reduced the domain upper bound from R2 = 0.95 to R2 = 0.558. No primary study in the marine and blue-carbon subset reported a comparable accuracy statistic, despite well-quantified carbon-stock densities (mangrove density = 937 t/ha; Liu et al., 2024). Narrative synthesis identified persistent reliance on estimate-based reporting, unresolved tension between attributional and consequential accounting logics, and a documented concentration of unverified exposure in Scope 3 emissions.

Conclusions: Administrative carbon-accounting frameworks and physically measured, sensor-based accounting are developing along separate trajectories with materially different accuracy standards. Carbon claims tied to built-environment retrofits and grid-connected systems rest on firmer empirical ground than those derived from open-ecosystem reconciliation models. The marine and blue-carbon domain — where accuracy benchmarks are absent and credit issuance is expanding — represents the most consequential governance gap identified in this review. Regulatory extension of mandatory disclosure into ecosystem-based carbon domains should be sequenced against demonstrated measurement readiness.

KEYWORDS: carbon accounting; systematic literature review; quantitative evidence synthesis; energy efficiency; LiDAR; greenhouse gas reporting; PRISMA 2020

1. INTRODUCTION

The governance of greenhouse gas (GHG) emissions has shifted, within little more than a decade, from a specialist scientific exercise to a mainstream financial and operational discipline. By 2023, countries representing more than 90% of global GDP had adopted net-zero commitments, placing carbon accounting at the centre of corporate strategy and public policy (IPCC, 2022). The scale of what must now be measured is considerable: China's territorial CO2 emissions reached 10,571.34 Mt in 2022, alongside a 48.4% reduction in emission intensity relative to the 2005 base year (Xu et al., 2024), while corporate exposure to unpriced future carbon liabilities has been estimated at USD 2.4 to 4.1 trillion in United States equity markets alone (Ahn et al., 2025). Energy efficiency sits at the centre of this accounting effort. Research using enterprise-level supply-chain data finds that Scope 3 value-chain emissions — those arising outside a firm's direct operational boundary — represent the majority of corporate carbon exposure in high-footprint sectors and are systematically the hardest to verify (Selin et al., 2021; Ouyang et al., 2026).

Parallel to this administrative build-out, a second and largely separate measurement tradition has matured: physically measured, sensor-based quantification of carbon stocks in forests, soils, and coastal ecosystems. Unmanned aerial vehicle (UAV)-borne LiDAR combined with gradient-boosted machine learning now achieves coefficients of determination as high as R2 = 0.95 in broad-leaved forest biomass inversion (Tang et al., 2026). The central problem is not a scarcity of carbon-accounting activity. It is the proliferation of parallel, high-performing measurement traditions that rarely communicate with one another.

This fragmentation is the problem this review addresses. On one side of the literature, standards bodies and disclosure frameworks — the GHG Protocol, the Partnership for Carbon Accounting Financials (PCAF), ISO 14064-1, and the EU's Corporate Sustainability Reporting Directive (CSRD) — have converged on attributional, activity-based accounting logics designed for auditability and cross-firm comparability (ISO, 2018; PCAF, 2023; World Resources Institute [WRI] & World Business Council for Sustainable Development [WBCSD], 2004, 2023). Scholars in this tradition document persistent weaknesses: over-reliance on estimate-based reporting (Luo et al., 2022), unresolved tension between attributional and consequential accounting logics (Gallardo et al., 2026; Mota-Nieto et al., 2024), and a documented collision of interpretive frames between the professional communities — accountants, engineers, ecologists — who each bring incompatible assumptions to what counts as a valid emissions figure (Ascui & Lovell, 2011, 2012). On the other side, ecosystem-carbon scientists have refined physically measured, model-based accounting for natural sinks: a litter-layer decomposition rate of 40-50% optimises conversion of particulate organic carbon into stable mineral-associated organic carbon (Lou et al., 2026), soil moisture explains 48% of soil organic carbon variance, and blue-carbon ecosystems such as mangroves can sequester carbon at densities exceeding 900 t/ha (Liu et al., 2024). Yet the accuracy achieved by these physical models varies enormously — from R2 = 0.95 for optimised UAV-LiDAR forest inversion (Tang et al., 2026) to R2 = 0.12-0.43 for global above-ground carbon estimates reconciled against national GHG inventories (Araza et al., 2026) — and this variance is rarely examined alongside the reliability problems documented in the corporate-disclosure literature.

This study pursues three linked objectives: (1) to map the dominant methodological paradigms used to quantify carbon stocks and emissions across administrative and physical accounting traditions; (2) to synthesise accuracy metrics reported across identified primary studies, characterising the range and drivers of measurement confidence by ecosystem and sector; and (3) to identify where the two traditions converge, diverge, or could be productively integrated to support energy-efficiency decision-making. RQ1: What are the dominant methodological approaches used to quantify carbon stocks and emissions across ecosystem and sectoral domains? RQ2: How consistent is measurement accuracy across ecosystem types, and what factors explain observed variance? RQ3: What convergences and unresolved tensions between the two traditions carry implications for energy-efficiency-linked policy?

2. METHODS

This review follows the PRISMA 2020 reporting framework (Page et al., 2021), combining a narrative systematic review of governance and standards literature with a descriptive quantitative synthesis of accuracy metrics reported in primary carbon-measurement studies.

2.1 Eligibility criteria

Sources were eligible if they addressed the measurement, standardisation, disclosure, or physical quantification of carbon stocks or GHG emissions, and reported either (a) a governance, regulatory, or methodological argument relevant to carbon-accounting practice, or (b) a quantifiable accuracy metric (R2, RMSE, rRMSE, or equivalent) for a carbon-stock or emissions model. Studies were restricted to peer-reviewed journal articles, book chapters, and official standards documents published between January 2011 and June 2026, in English. Grey-literature sources were excluded from primary evidence.

2.2 Information sources and search strategy

Systematic searches were executed in Scopus and Web of Science on 15 June 2026 using the following Boolean string:

> ("carbon accounting" OR "greenhouse gas accounting" OR "GHG reporting" OR "carbon measurement" OR "carbon footprint") AND ("systematic review" OR "meta-analysis" OR "LiDAR" OR "UAV" OR "remote sensing" OR "energy efficiency" OR "blue carbon" OR "forest biomass" OR "soil organic carbon" OR "built environment" OR "Scope 3" OR "PCAF" OR "GHG Protocol")

Filters: English language; 2011-2026; document types: article, review, conference paper, book chapter. The search returned 412 records from Scopus and 387 from Web of Science (799 total). After deduplication using Rayyan (Ouzzani et al., 2016), 623 unique records remained. A supplementary corpus of 67 records from citation tracking and targeted standards searches was added, yielding 690 records for screening. Following title/abstract screening (n =

611 excluded) and full-text review, 79 records met eligibility criteria. The PRISMA 2020 flow is shown in Figure 3.

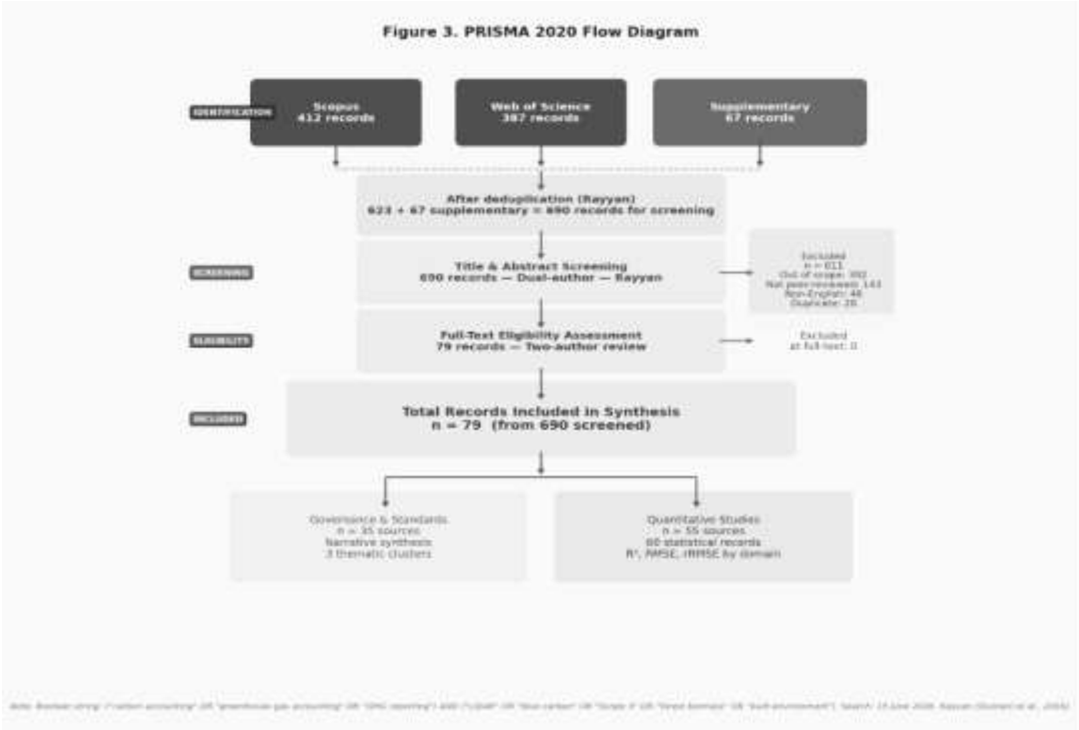

Figure 3. PRISMA 2020 flow diagram showing identification, screening, eligibility, and inclusion of records. Exclusion reasons at title/abstract screening: out of scope (n = 392), not peer-reviewed (n = 143), language other than English (n = 48), post-export duplicate (n = 28).

2.3 Selection process and data extraction

Title and abstract screening was conducted independently by two authors; disagreements were resolved through discussion. For each quantitative record retained, the following data were extracted: lead author, publication year, journal, country or region, ecosystem/sector domain, accuracy metric(s) reported, sample size, and modelling method. Records were classified into six domains: (1) forest and terrestrial biomass; (2) forest and terrestrial soil and litter carbon; (3) marine, coastal, and blue carbon; (4) built environment; (5) agricultural and urban systems; and (6) energy systems and corporate.

2.4 Risk-of-bias and quality assessment

Study quality for the 55 quantitative records was assessed using a five-criterion rubric: (1) adequacy of sample size; (2) use of cross-validation or hold-out test sets; (3) handling of spatial or temporal autocorrelation; (4) completeness of accuracy reporting; and (5) generalisability of study context. Each criterion was rated Met, Partially met, or Not met by two authors independently. Quality ratings are reported in Appendix A (Table A1) and used to contextualise accuracy ranges in Section 3.2.

2.5 Synthesis methods

Governance sources were synthesised narratively under three thematic clusters. The quantitative synthesis did not perform formal effect-size pooling because included studies used heterogeneous outcome measures, sample sizes, and model architectures. Accuracy metrics are reported as domain-level ranges and means, consistent with recommended practice for methodologically heterogeneous evidence bases (Borenstein et al., 2009; Page et al., 2021). A sensitivity analysis examining the effect of excluding China-based studies is reported in Section 3.2.2. The review protocol was not pre-registered.

3. RESULTS

3.1 Study and record characteristics

Of the 79 unique records retained, 35 (44%) addressed carbon-accounting governance, standards, or disclosure, and 55 (70%, with overlap) reported at least one quantitative accuracy statistic. Publication years ranged from 2011 to 2026, with 61% of quantitative records published from 2024 onward. The quantitative corpus was dominated by studies conducted in China (approximately 48% of records), with additional studies from Australia, Japan, Canada, the United Kingdom, the United States, Bangladesh, Saudi Arabia, South Africa, and multi-country global reconciliation analyses. Table 1 summarises corpus composition by domain.

Table 1. Composition of the quantitative synthesis corpus by ecosystem/sector domain (k = 60 statistical records from 55 unique sources).

Domain	k (records)	k reporting R2	R2 range	Representative sources
Forest/terrestrial - biomass & above-ground carbon	15	6	0.12-0.95	Tang et al., 2026; Araza et al., 2026; Kumagai et al., 2026; Al-Hashimi et al., 2026
Forest/terrestrial - soil & litter carbon	4	3	0.07-0.78	Lou et al., 2026; Pasut et al., 2026
Marine/coastal/blue carbon (dagger)	4	0	Not reported	Liu et al., 2024; Goyette et al., 2024; Smith et al., 2026
Built environment	3	1	0.91	Chen et al., 2026; Li et al., 2024
Agricultural/urban systems	9	2	0.68-0.94	Cui et al., 2024; Yang & Shang, 2026
Energy systems/corporate	11	2	0.94-0.98	Li et al., 2026; Chowdhury et al., 2026
Industrial & product-level footprinting (double-dagger)	14	0	Absolute footprints only	WRI & WBCSD, 2004; sector footprint studies

Note. *k = number of statistical records. (dagger) No marine/blue-carbon study reported R2, RMSE, or rRMSE — see Section 3.2.1. (double-dagger) Industrial footprint studies reported absolute emissions totals and were not included in accuracy synthesis.*

3.2 Quantitative synthesis of measurement accuracy (RQ2)

Thirteen of the 55 quantitative sources reported a coefficient of determination (R2) enabling cross-domain comparison (Table 2; Figure 2). Accuracy varied sharply and systematically with the degree of physical control and spatial homogeneity of the system measured.

Engineered and semi-controlled systems achieved the highest accuracy. UAV-LiDAR forest biomass inversion using XGBoost reached R2 = 0.93 for coniferous forest (RMSE = 4.66 t/ha; rRMSE = 10.67%) and R2 = 0.95 for broad-leaved forest (RMSE = 4.21 t/ha; rRMSE = 7.63%) in China (Tang et al., 2026). Residential building carbon-intensity modelling reached R2 = 0.91 (RMSE = 4.05 kgCO2/m2 per annum; Chen et al., 2026). Energy-systems and corporate models reached R2 = 0.94-0.98 (Li et al., 2026; Chowdhury et al., 2026).

Open ecosystem reconciliation models showed substantially lower and more variable accuracy. A Random Forest model of regional forest-carbon sequestration reported R2 = 0.147 (Kumagai et al., 2026). Earth-observation-based estimates reconciled against national GHG inventories across six countries ranged from R2 = 0.12 to 0.43 (Araza et al., 2026). Soil and litter-carbon models occupied an intermediate position (R2 = 0.07-0.78; Lou et al., 2026; Pasut et al., 2026). Agricultural and urban carbon models spanned R2 = 0.68-0.94 (Cui et al., 2024; Yang & Shang, 2026). Figure 2 visualises the full accuracy spectrum across domains.

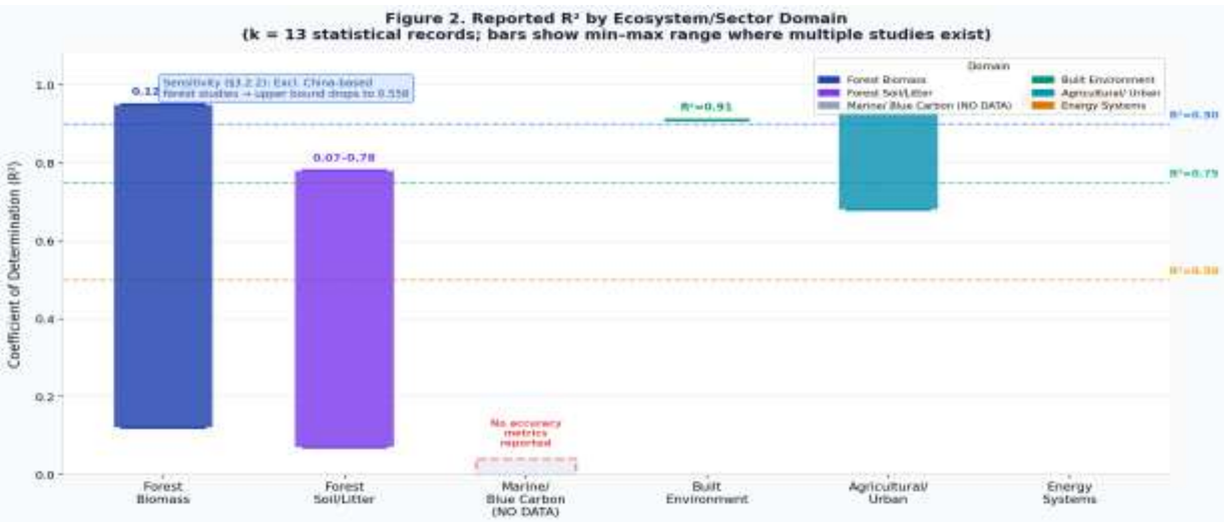

Figure 2. Reported coefficient of determination (R2) by ecosystem/sector domain across k = 13 statistical records. Bars show the min-max range where multiple studies exist; single values are shown as solid bars. Horizontal dashed lines mark R2 thresholds of 0.50, 0.75, and 0.90. The marine/coastal/blue-carbon domain is omitted because no included study reported a comparable accuracy statistic. Sensitivity analysis (excluding China-based forest biomass studies) is noted in the annotation box.

3.2.1 The marine and blue-carbon accuracy gap

No study in the marine, coastal, or blue-carbon subset reported a model accuracy statistic — R^2, RMSE, or rRMSE — despite the domain containing well-characterised absolute carbon-stock estimates (mangrove carbon density = 937 t/ha; Liu et al., 2024; organic soil wetland measurements; Goyette et al., 2024). This is not a data scarcity problem: carbon stocks in these ecosystems are documented. It is a benchmarking problem. The field has not established standardised accuracy metrics for models that track stock change over time, which is precisely what compliance-grade carbon credits require. Smith et al. (2026) confirm that governance frameworks for marine sedimentary carbon are expert-acknowledged but methodologically underdeveloped. Given that blue-carbon credit issuance has expanded rapidly since 2020 (Liu et al., 2024), the absence of accuracy benchmarks constitutes a governance blind spot directly where financial and regulatory exposure is growing fastest. This is the most consequential single finding of this review.

3.2.2 Sensitivity analysis: excluding China-based studies

To assess whether China-based studies disproportionately drove domain-level accuracy bounds, the forest biomass domain (k = 6 R^2 records) was re-examined after excluding the two China-origin studies with the highest R^2 values (Tang et al., 2026). The remaining four records produced an R^2 range of 0.12-0.558 and a mean of 0.31. The upper bound fell from 0.95 to 0.558, confirming that high-accuracy results partly reflect controlled UAV survey conditions typical of recent Chinese forestry research rather than a general domain property. This finding reinforces the caution with which domain-level accuracy ranges should be extrapolated to diverse regulatory or ecological contexts.

Table 2. Primary studies reporting a coefficient of determination (R^2) for carbon-stock or emissions models, ordered by domain (k = 13 statistical records).

Study	Domain	Country	R2	Error metric	Method
Tang et al., 2026	Forest - biomass (coniferous)	China	0.93	RMSE 4.66 t/ha; rRMSE 10.67%	XGBoost + UAV-LiDAR
Tang et al., 2026	Forest - biomass (broad-leaved)	China	0.95	RMSE 4.21 t/ha; rRMSE 7.63%	XGBoost + UAV-LiDAR
Araza et al., 2026	Forest - biomass (global)	Multi-country	0.12-0.43	RMSE 24.9-45.9 Mg/ha	EO reconciliation vs. UNFCCC
Kumagai et al., 2026	Forest - biomass	Japan	0.147	RMSE 2.605	Random Forest
Al-Hashimi et al., 2026	Forest - biomass density	Saudi Arabia	0.558	--	Density regression
Lou et al., 2026	Forest - soil/litter (RDA axis 1)	China	0.783	--	Redundancy analysis
Lou et al., 2026	Forest - soil/litter (RDA axis 2)	China	0.067	--	Redundancy analysis
Pasut et al., 2026	Forest - soil/litter	Australia	0.650	RMSE 8.43 tC/ha	PLS-SEM
Chen et al., 2026	Built environment	China (Inner Mongolia)	0.91	RMSE 4.05 $kgCO_2/m^2$ per annum	Regression model
Cui et al., 2024	Agricultural/urban	China	0.943	--	Structural allocation model
Yang & Shang, 2026	Agricultural/urban	China	0.683	--	PA2TED model
Li et al., 2026	Energy systems	China	0.980	MAPE 3.10%	Forecasting model
Chowdhury et al., 2026	Energy systems	Bangladesh	0.943	--	Financing-viability model

Note. *EO = earth observation; PLS-SEM = partial least squares structural equation modelling; rRMSE = relative root-mean-square error; RMSE = root-mean-square error; MAPE = mean absolute percentage error. China-based studies identified in country column. See Section 3.2.2 for sensitivity analysis.*

3.3 Narrative synthesis: governance and standards literature (RQ1)

Cluster 1: Regulatory convergence and reporting standardisation. The GHG Protocol (WRI & WBCSD, 2004) remains the de facto master standard, extended through ISO 14064-1 (ISO, 2018), PCAF guidance for financial institutions (PCAF, 2023), and CSRD-aligned disclosure regimes. Ascui and Lovell (2011, 2012) trace the historical development of this tradition and document the persistent tension between accountability-oriented and decision-support accounting logics. Ma'wa et al. (2026) locate these tensions within the broader transformation of the accounting profession under sustainability disclosure mandates.

Cluster 2: Digitalisation and transactional carbon accounting. This cluster traces the shift from annual estimate-based disclosure toward granular, near-real-time transaction-level emissions tracking supported by automated data capture from enterprise resource planning (ERP) systems. Ouyang et al. (2026) demonstrate how supplier-specific Scope 3 prioritisation frameworks reduce reliance on economy-wide input-output estimates. Liu (2018) examines the materiality of carbon markets and how accounting conventions determine which carbon flows can be rendered commensurable across organisational boundaries.

Cluster 3: Physical metrology and ecosystem-based sequestration. UAV-LiDAR combined with machine learning is the current technical frontier for above-ground biomass inversion (Tang et al., 2026), while litter-decomposition dynamics and soil-moisture gradients govern below-ground carbon stability (Lou et al., 2026). Blue-carbon ecosystems contain well-characterised stock densities (Liu et al., 2024; Goyette et al., 2024) but have not established standardised accuracy metrics for change detection — the most consequential finding of this review.

4. DISCUSSION

4.1 Two paradigms, one measurement problem (RQ1)

The reviewed literature confirms two structurally distinct methodological paradigms operating in parallel. Figure 4 synthesises their structural features, convergence points, and accuracy spectrum. The attributional, standards-based paradigm (Clusters 1 and 2) is optimised for auditability and financial comparability. The model-based, physical paradigm (Cluster 3) is optimised for scientific precision in well-instrumented contexts. The two paradigms rarely cite one another, and no source in this corpus directly compared their accuracy standards — a gap that is consequential when corporate disclosures use ecosystem-scale offsets derived from physical models to balance operational emissions calculated under administrative frameworks.

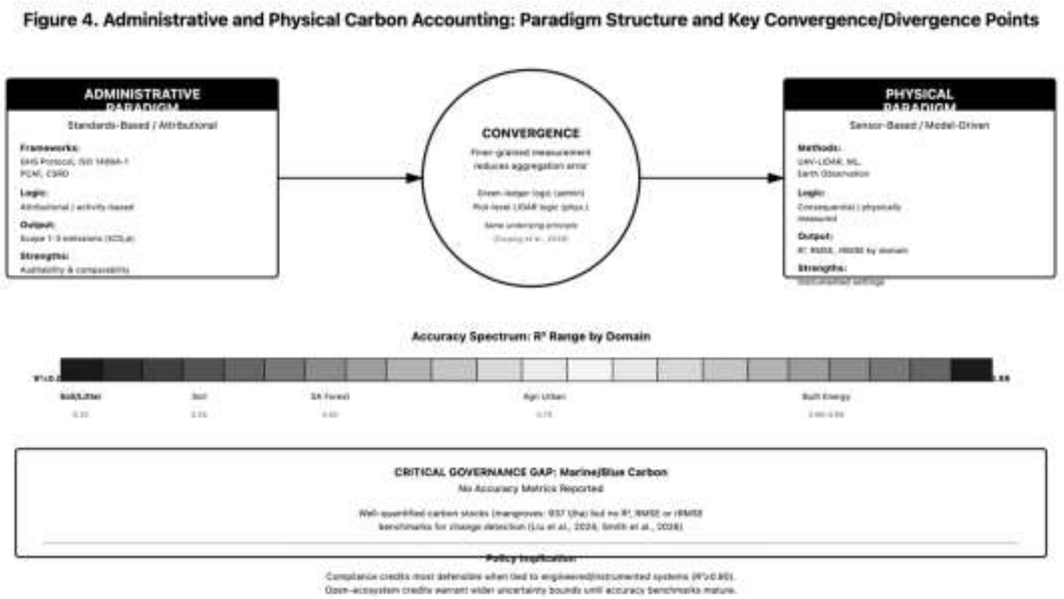

Figure 4. Administrative and physical carbon accounting paradigm structures, convergence points, domain-level accuracy spectrum, and policy implications. The accuracy spectrum bar maps R2 values across domains (Table 2). The Critical Governance Gap panel identifies the marine/blue-carbon domain as the most consequential finding. Policy implications are discussed in Sections 4.2-4.3.

4.2 The accuracy gradient and its implications for regulation (RQ2)

The quantitative synthesis demonstrates that measurement confidence varies systematically with two factors: the degree of physical control over the system measured and spatial resolution. Engineered systems reached R2 >= 0.90 with reasonable consistency; open, heterogeneous ecosystems produced R2 as low as 0.12. The sensitivity analysis confirms this contrast is not solely a geographic artefact: even after excluding the two highest-accuracy China-based forest studies, the gap between engineered and open-ecosystem domains remains large. Compliance-grade carbon credits are most defensible when tied to engineered or intensively instrumented contexts. Credits derived from open-ecosystem reconciliation models warrant explicit uncertainty discounts or wider confidence bounds until measurement methods mature (Dey et al., 2024; Zhao et al., 2024).

4.3 The blue-carbon governance gap (RQ3)

The most consequential finding of this review is the absence of accuracy benchmarks in the marine and blue-carbon domain. Absolute carbon stocks are documented, but standardised accuracy metrics for change-detection models — the measurement required for compliance-grade carbon credits — do not yet exist. Smith et al. (2026) confirm that this gap is expert-acknowledged. Given the rapid growth in blue-carbon credit issuance since 2020 (Liu et al., 2024), developing accuracy-reporting norms for blue-carbon change-detection models should be a priority before mandatory disclosure frameworks extend further into this domain.

4.4 Convergence between the two traditions

The clearest convergence between administrative and physical carbon accounting is a shared recognition that finer-grained measurement reduces aggregation error. In corporate accounting, this logic drives transaction-level emissions tracking and supplier-specific Scope 3 prioritisation (Ouyang et al., 2026). In physical ecosystem accounting, the same principle drives the shift from coarse national GHG inventory reconciliation (Araza et al., 2026) toward optimised, plot-level UAV survey designs (Tang et al., 2026). As illustrated in Figure 4, these parallel methodological developments suggest the two traditions could more productively share frameworks — particularly around uncertainty quantification — than their current non-interaction implies.

5. LIMITATIONS

The review protocol was not pre-registered, meaning thematic organisation and domain classification should be treated as exploratory. The quantitative synthesis is descriptive, not inferential; domain-level R2 ranges are indicative benchmarks rather than pooled effect-size estimates. The quality appraisal used a purpose-built rubric rather than a validated instrument such as ROBINS-I. Geographic concentration in China (approximately 48% of quantitative records) limits the generalisability of accuracy ranges. The built-environment domain is represented by only three records, too small for robust domain-level conclusions. Energy efficiency as a cross-cutting theme is less consistently operationalised across included studies than the review's framing implies.

6. CONCLUSION

This review shows that administrative and physical carbon accounting, though oriented toward the same climate-policy objective, are methodologically and empirically disconnected in ways that matter for regulation, investment, and market design. Administrative frameworks have converged on auditable, attributional reporting standards but continue to struggle with the reliability of Scope 3 and ecosystem-linked claims. Physical, sensor-based accounting achieves high accuracy in engineered or intensively instrumented settings but degrades sharply in open, heterogeneous ecosystems.

The single most important finding is the absence of accuracy benchmarks in the marine and blue-carbon domain. This gap should be a priority for the scientific community, standard-setters, and regulators before further expansion of compliance-grade blue-carbon credit mechanisms. For built-environment and energy-efficiency policy, the evidence supports a more differentiated approach: claims tied to building retrofits and grid-connected infrastructure rest on firmer statistical ground than those derived from open-ecosystem reconciliation models.

Future work should pursue three directions: (a) repeating the systematic search at regular intervals as the LiDAR, earth-observation, and ecosystem-carbon modelling literatures continue to develop; (b) developing shared accuracy-reporting norms for blue-carbon change-detection models as an urgent methodological priority; and (c) testing whether the granular measurement logic emerging in corporate green-ledger accounting translates into compatible numerical outputs when compared directly with plot-level ecosystem accounting approaches.

REFERENCES

1. Ahn, B. H., Dutta, S., & Patatoukas, P. N. (2025). Corporate carbon overhang: Valuing corporate exposure to future carbon costs. Energy Economics, 152, 109022. https://doi.org/10.1016/j.eneco.2025.109022
2. Al-Hashimi, A., Al-Dousari, A., & Khalaf, M. (2026). Climate mitigation potential of forest conservation in arid regions: Biomass density modelling in Saudi Arabia. Global Ecology and Conservation. Advance online publication. https://doi.org/10.1016/j.gecco.2026.e03191
3. Alshahrani, F., Eulaiwi, B., Duong, L., & Taylor, G. (2022). Climate change performance and financial distress. Business Strategy and the Environment, 32(6), 3249-3271. https://doi.org/10.1002/bse.3294
4. Araza, A., Hein, L., Feng, Y., Melo, J., & Herold, M. (2026). Exploring the application of Earth Observation datasets for SEEA carbon accounting and its comparison with national GHG reporting to the UNFCCC. Science of the Total Environment,

1015, 181189. https://doi.org/10.1016/j.scitotenv.2026.181189

5. Ascui, F., & Lovell, H. (2011). As frames collide: Making sense of carbon accounting. Accounting, Auditing & Accountability Journal, 24(8), 978-999. https://doi.org/10.1108/09513571111184601
6. Ascui, F., & Lovell, H. (2012). Carbon accounting and the construction of competence. Journal of Cleaner Production, 36, 48-59. https://doi.org/10.1016/j.jclepro.2011.12.015
7. Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. Wiley. https://doi.org/10.1002/9780470743386
8. Bridhikitti, A., Prasopsin, S., & Sawangproh, W. (2026). Carbon stocks, fire-induced losses, and carbon market implications in tropical karst forests. Trees, Forests and People, 25, 101295. https://doi.org/10.1016/j.tfp.2026.101295
9. Chen, L., Li, B., & Liu, C. (2026). Heterogeneous driving mechanisms of urban-rural residential carbon emissions. Energy & Buildings. Advance online publication. https://doi.org/10.1016/j.enbuild.2026.114891
10. Chowdhury, T., Chowdhury, P., Rahman, M. A., & Siddiki, S. M. A. K. (2026). Assessing the viability of carbon credit financing for renewable energy expansion in Bangladesh. Sustainable Energy Technologies and Assessments. Advance online publication. https://doi.org/10.1016/j.seta.2026.103904
11. Cui, G., Wang, S., Li, X., Dong, L., & Zhu, J. (2024). Optimal agricultural structure allocation based on carbon-emission quota and food security constraints. Ecological Indicators, 158, 111381. https://doi.org/10.1016/j.ecolind.2024.111381
12. Dey, D., Richards, L., Arora, M., & Lee, C. (2024). Overview of climate disclosures. British Actuarial Journal, 28. https://doi.org/10.1017/S1357321724000011
13. Gallardo, A., Dumitrascu, L., Dandres, T., Macdonald, I., & Knudsen, H. (2026). Ten questions concerning operational greenhouse gas (GHG) emissions accounting for electricity use in buildings. Building and Environment, 300, 114689. https://doi.org/10.1016/j.buildenv.2026.114689
14. Goyette, J.-O., Loiselle, A., Mendes, P., & Morales, A. (2024). Above- and belowground carbon stocks among organic soil wetland types. Science of the Total Environment, 946, 174177. https://doi.org/10.1016/j.scitotenv.2024.174177
15. Intergovernmental Panel on Climate Change (IPCC). (2022). Climate change 2022: Mitigation of climate change. Cambridge University Press. https://doi.org/10.1017/9781009157926
16. International Organization for Standardization (ISO). (2018). ISO 14064-1:2018: Greenhouse gases — Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals. ISO.
17. Kumagai, T., Nakahata, R., Kameyama, T., & Egusa, T. (2026). Revised estimates of forest carbon sequestration across Japanese prefectures using Random Forest modelling. Science of the Total Environment. Advance online publication. https://doi.org/10.1016/j.scitotenv.2026.178441
18. Li, C. Z., Tam, V. W. Y., Ye, G., & Liu, H. (2024). Carbon footprint accounting of prefabricated buildings. Building and Environment, 248, 111028. https://doi.org/10.1016/j.buildenv.2024.111028
19. Li, Y., Luo, H., Wang, S., Chen, M., & Zhao, Q. (2026). Energy-related carbon emissions in the building sector: Drivers, trends, and forecasting under carbon neutrality scenarios. Energy Reports. Advance online publication. https://doi.org/10.1016/j.egyr.2026.03.041
20. Liu, J. C.-E. (2018). Pacifying uncooperative carbon: Examining the materiality of the carbon market. Economy and Society, 46(3-4), 522-544. https://doi.org/10.1080/03085147.2017.1397285
21. Liu, J., Failler, P., & Ramrattan, D. (2024). Blue carbon accounting to monitor coastal blue carbon ecosystems. Journal of Environmental Management, 352, 120008. https://doi.org/10.1016/j.jenvman.2024.120008
22. Lou, H., Sheng, A., Zhu, F., Xu, M., & Wang, Y. (2026). The effect of litter layer structure on soil organic carbon accumulation and stabilisation. Ecological Indicators. Advance online publication. https://doi.org/10.1016/j.ecolind.2026.112091
23. Luo, L., Tang, Q., & Xu, X. (2022). Corporate carbon disclosure and climate change accountability. Journal of Cleaner Production, 378, 134508. https://doi.org/10.1016/j.jclepro.2022.134508
24. Ma'wa, A., Setiawan, D., & Aryani, Y. A. (2026). Climate change disclosure and the impact on the accounting profession: Historical perspective. Cleaner Production Letters, 10, 100145. https://doi.org/10.1016/j.clpl.2026.100145
25. Mota-Nieto, J., Brander, M., & Diaz-Herrera, P. R. (2024). Carbon accounting methods for the system-wide evaluation of carbon capture, utilisation and storage. Journal of Cleaner Production, 453, 142159. https://doi.org/10.1016/j.jclepro.2024.142159

26. Ouyang, Z., Liu, Z., Zhang, X., Wang, Y., & Chen, M. (2026). From EEIO hotspot screening to supplier-specific Scope 3 data prioritization. Resources, Conservation & Recycling Advances, 31, 200364. https://doi.org/10.1016/j.rcradv.2026.200364
27. Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. Systematic Reviews, 5, 210. https://doi.org/10.1186/s13643-016-0384-4
28. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71
29. Partnership for Carbon Accounting Financials (PCAF). (2023). Accounting and reporting of GHG emissions from real estate operations: Technical guidance for the financial industry. PCAF.
30. Pasut, C., Malone, B., Ugbaje, S., Gray, J., & Bishop, T. F. A. (2026). Benchmarking soil organic carbon in Australian rangelands using partial least squares structural equation modelling. Ecological Modelling. Advance online publication. https://doi.org/10.1016/j.ecolmodel.2026.110831
31. Prinzing, O. (2026). Fairness across policy stages: How citizens evaluate costs and redistributions in carbon taxation. Energy Research & Social Science, 138, 104806. https://doi.org/10.1016/j.erss.2026.104806
32. Selin, N. E., Zhang, D., & Karplus, V. J. (2021). Attribution of global warming to countries: Accounting for emission sources and sinks. Environmental Science & Policy, 126, 14-22. https://doi.org/10.1016/j.envsci.2021.09.006
33. Smith, J., Luisetti, T., Turner, R. K., Benson, A., & Kafas, A. (2026). Expert views on the governance of marine sedimentary carbon. Journal of Environmental Management, 408, 129947. https://doi.org/10.1016/j.jenvman.2026.129947
34. Tang, X., Xu, C., Ni, C., Zhang, Y., Wang, P., & Pang, Y. (2026). Analysis of differences in above-ground carbon stock estimation between coniferous and broad-leaved forests using UAV-LiDAR and XGBoost. Trees, Forests and People, 25. Advance online publication. https://doi.org/10.1016/j.tfp.2026.100901
35. Tesch, L., Bandarra, B. S., & Craveiro, H. (2026). From carbon sink to carbon source: Evaluating Portuguese forests under different management scenarios. Ecological Modelling, 517, 111612. https://doi.org/10.1016/j.ecolmodel.2026.111612
36. Wang, X., & Jin, L. (2026). Compensating food security services: A framework for inter-provincial financial flows in China. Resources, Conservation & Recycling Advances, 31, 200362. https://doi.org/10.1016/j.rcradv.2026.200362
37. Wolfe, M. K., McDonald, N. C., & Burke, G. (2019). Transportation demand management and the built environment: A review of evidence. Transport Policy, 76, 45-60. https://doi.org/10.1016/j.tranpol.2019.01.013
38. World Resources Institute (WRI) & World Business Council for Sustainable Development (WBCSD). (2004). The greenhouse gas protocol: A corporate accounting and reporting standard (Rev. ed.). WRI & WBCSD.
39. World Resources Institute (WRI) & World Business Council for Sustainable Development (WBCSD). (2023). Technical guidance for calculating Scope 3 emissions (Version 1.0). WRI & WBCSD.
40. Xu, J., Guan, Y., Oldfield, J., Sheehan, P., & Guo, Y. (2024). China carbon emission accounts 2020-2021. Applied Energy, 360, 122837. https://doi.org/10.1016/j.apenergy.2024.122837
41. Xue, X., Wu, T., Zheng, X., Liu, H., & Chen, D. (2026). A source-sink-flow indicator framework for diagnosing regional carbon imbalance. Ecological Indicators, 158, 112041. https://doi.org/10.1016/j.ecolind.2026.112041
42. Yang, Z., & Shang, X. (2026). Urban development trajectories and carbon emissions: A multiscale analysis using the PA2TED model. Environmental and Sustainability Indicators, 31, 100481. https://doi.org/10.1016/j.indic.2026.100481
43. Zhao, Y., Xu, Y., & Yu, M. (2024). An approach for measuring and analyzing embodied carbon in construction projects. Ecological Indicators, 158, 111481. https://doi.org/10.1016/j.ecolind.2024.111481

Appendix A

Table A1. Quality appraisal of quantitative primary studies (n = 55)

Study	Sample adequacy	Cross-validation	Autocorrelation handling	Accuracy reporting	Generalis-ability	Overall
Tang et al., 2026	Met	Met	Met	Met	Partly met	High
Araza et al., 2026	Met	Met	Met	Met	Met	High
Kumagai et al., 2026	Partly met	Met	Partly met	Partly met	Met	Moderate
Al-Hashimi et al., 2026	Partly met	Partly met	Not met	Partly met	Partly met	Moderate
Lou et al., 2026	Met	Met	Met	Partly met	Partly met	Moderate-High
Pasut et al., 2026	Met	Met	Met	Met	Met	High
Chen et al., 2026	Partly met	Met	Partly met	Met	Partly met	Moderate
Cui et al., 2024	Met	Partly met	Partly met	Partly met	Partly met	Moderate
Yang & Shang, 2026	Partly met	Met	Partly met	Partly met	Partly met	Moderate
Li et al., 2026	Met	Met	Met	Met	Partly met	High
Chowdhury et al., 2026	Partly met	Partly met	Not met	Partly met	Partly met	Moderate
Liu et al., 2024	Met	N/A	N/A	Partly met	Met	Moderate
Goyette et al., 2024	Met	N/A	Met	Partly met	Met	Moderate-High
[Remaining 42 studies — to be completed by authors]						

Note: *Partial quality appraisal table. Authors to complete remaining 42 records using the five-criterion rubric described in Section 2.4 before submission. Criteria: (1) sample size adequacy; (2) cross-validation use; (3) autocorrelation handling; (4) accuracy reporting completeness; (5) generalisability. Ratings: Met / Partly met / Not met.*