

A GRI-BASED ESG DISCLOSURE DATASET FOR PALM OIL, PULP AND PAPER, AND FORESTRY COMPANIES IN INDONESIA

Tegar Ditya Pragama^{1*}, Disman², Heny Hendrayati³, Puspo Dewi Dirgantari⁴

^{1,2,3,4} Universitas Pendidikan Indonesia

E-mail: tegarditya@upi.edu^{1*}, disman@upi.edu², henyhendrayati@upi.edu³, puspodewi@upi.edu⁴

Received: 01/04/2026 | Revised: 05/04/2026 | Accepted: 20/04/2026 | Published: 18/05/2026

Abstract

This article presents an original Environmental, Social, and Governance (ESG) disclosure dataset constructed through systematic dual-coder content analysis for Indonesian forest-utilizing companies listed on the Indonesia Stock Exchange (IDX). The dataset encompasses 71 company-year observations from 41 unique companies across fiscal years 2023 and 2024, covering three sectors: palm oil (21 companies, 37 observations), pulp and paper (11 companies, 21 observations), and forestry and wood products (9 companies, 13 observations). ESG disclosure was assessed using 144 indicators derived from the Global Reporting Initiative (GRI) Standards, comprising 72 Governance, 32 Environmental, and 40 Social items. Each indicator was scored using a binary approach (1 = disclosed, 0 = not disclosed), with inter-coder reliability yielding Cohen's kappa coefficients of 0.85 to 0.91. This dataset addresses a critical gap in ESG disclosure research for Indonesia's forest-utilizing sector, which faces increasing international scrutiny under the European Union Deforestation Regulation (EUDR). The data can be reused for analyzing disclosure determinants, conducting subsector comparisons, evaluating the effectiveness of sustainability certification schemes, and benchmarking corporate transparency against regulatory requirements. The full dataset is publicly deposited at Mendeley Data (DOI: 10.17632/hnsb3h9353.2) under a CC BY 4.0 license.

Keywords: ESG disclosure; GRI Standards; Sustainability reporting; Indonesia Stock Exchange; Forest-utilizing sector; Binary scoring; Palm oil; EUDR

INTRODUCTION

The assessment of Environmental, Social, and Governance (ESG) disclosure practices has gained increasing importance as stakeholders demand greater corporate transparency on sustainability issues. The Global Reporting Initiative (GRI) Standards represent the most widely adopted framework for sustainability reporting worldwide, providing a comprehensive set of indicators across governance, environmental, and social dimensions [1]. In Indonesia, the Financial Services Authority (Otoritas Jasa Keuangan, OJK) introduced Regulation No. 51/POJK.03/2017, mandating financial institutions and publicly listed companies to publish sustainability reports [2]. Indonesia's forest-utilizing sector holds particular significance in this context: the country is the world's largest palm oil producer, accounting for approximately 54% of global exports [3], while its pulp and paper industry ranks among the largest in Asia. Oil palm plantations alone cover approximately 16.4 million hectares [4], and the sector directly and indirectly employs over 16 million people [3]. Given the scale of these industries and their environmental footprint, understanding ESG disclosure practices within Indonesia's forest-utilizing companies is both economically and ecologically critical.

Despite this significance, transparency in Indonesia's forest-commodity sectors remains a concern. The European Union Deforestation Regulation (EUDR), which entered into force in June 2023, requires that commodities including palm oil, wood, and pulp products entering the EU market must be verified as deforestation-free, with full application for large operators postponed to 30 December 2026 [5]. This regulation places direct pressure on Indonesian exporters to demonstrate compliance through enhanced ESG disclosure. Meanwhile, assessments by the Zoological Society of London's SPOTT initiative indicate that the average ESG transparency score of assessed palm oil companies was only 48.2% in 2024, with non-RSPO members scoring as low as 17.4% [6]. Existing ESG disclosure studies have predominantly focused on broad market indices, such as Phung et al. [7] who assessed 20 VNSI-listed companies in Vietnam and Rosman et al. [8] who examined the top 100 Malaysian companies. Within Indonesia, Gutierrez-Ponce and Wibowo [9] analyzed SDG disclosure among 25

SRI-KEHATI index companies, while Soleha and Rosiana [10] explored sustainability reporting in the Indonesian palm oil industry. However, no publicly available dataset systematically documents ESG disclosure at the indicator level for Indonesia's forest-utilizing sector, and the pulp and paper subsector remains virtually unstudied.

To address this gap, this article presents a comprehensive ESG disclosure dataset for 41 Indonesian forest-utilizing companies listed on the Indonesia Stock Exchange (IDX), covering 71 company-year observations across fiscal years 2023 and 2024. The dataset spans three sectors: palm oil (21 companies, 37 observations), pulp and paper (11 companies, 21 observations), and forestry and wood products (9 companies, 13 observations). ESG disclosure was assessed using 144 indicators derived from the GRI Standards, structured into three dimensions: Governance (72 items encompassing GRI 102-103 and GRI 201-206), Environmental (32 items from GRI 301-308), and Social (40 items from GRI 401-419). Each indicator was scored using a binary approach (1 = disclosed, 0 = not disclosed) through systematic content analysis of annual reports and sustainability reports, consistent with established methodologies [7,8]. To the best of the authors' knowledge, this represents the first publicly available, GRI-based, indicator-level ESG disclosure dataset for Indonesia's forest-utilizing sector.

This dataset was developed as part of a doctoral research program at Universitas Pendidikan Indonesia (UPI). A separate empirical study by the present authors using a different measurement instrument—the Strategic ESG Capability Index (SECI), a 54-indicator composite framework—has been accepted for publication in *JDM* (*Jurnal Dinamika Manajemen*) [13]. That study tests mediation hypotheses through PLS-SEM on a broader cross-sector sample of Indonesian natural resource companies. The two workstreams employ distinct measurement instruments (binary GRI 144-indicator vs. SECI 54-indicator composite), distinct samples (forest-utilizing focus vs. broader cross-sector), and pursue distinct research outputs (open data publication vs. empirical hypothesis testing); the granular indicator-level data described in the present article are not contained in [13].

METHOD

The dataset presented in this article was constructed through systematic content analysis of publicly available corporate reports. While the source documents constitute publicly available publications, the dataset itself is an original scholarly product. First, the 144-indicator GRI assessment framework constitutes a unique variable set that, to the best of the authors' knowledge, is not available from any existing public source for the Indonesian forest-utilizing sector. Second, the scientific protocol for data construction, including dual-coder content analysis methodology, binary scoring criteria, and inter-coder reliability testing yielding Cohen's kappa coefficients of 0.85 to 0.91, is fully documented in this article. Third, the systematic transformation of hundreds of pages of unstructured corporate reports into a structured, quantitative, machine-readable dataset adds significant scholarly value beyond what exists in the primary sources. A complete list of all primary data sources with document types and access URLs is provided in Supplementary Table S1.

Sample Selection

The sample was drawn from companies listed on the Indonesia Stock Exchange (IDX) operating in forest-utilizing sectors. Three sectors were identified based on their direct linkage to forest-risk commodities regulated under the EUDR: (a) palm oil plantation and processing, (b) pulp and paper manufacturing, and (c) forestry and wood products. Companies were identified through IDX sectoral classifications and cross-referenced with industry descriptions in annual reports to verify primary business activities.

The inclusion criteria required companies to (1) be listed on IDX during the fiscal years 2023 or 2024, (2) derive a substantial portion of revenue from forest-utilizing activities, and (3) have publicly available annual reports or sustainability reports for the relevant fiscal year. Companies that were suspended from trading, underwent delisting during the observation period, or lacked publicly accessible reports were excluded. The initial population comprised 48 forest-utilizing companies. After applying the inclusion criteria and removing observations with insufficient report availability, the final sample consisted of 41 unique companies yielding 71 company-year observations. The extended dataset, deposited in the Mendeley repository, additionally includes observations from four non-forest sectors (packaging, pharmaceutical, food processing, and rubber), totaling 108 observations from 63 companies across seven sectors.

Data Sources and Collection Period

ESG disclosure data were collected from two primary document types: annual reports (*laporan tahunan*) and sustainability reports (*laporan keberlanjutan*). Reports were downloaded from the IDX website (www.idx.co.id) and individual company websites between August and December 2025. When a company

published both an annual report and a standalone sustainability report, the sustainability report was used as the primary source for ESG assessment, as these documents typically provide more detailed disclosure on environmental and social matters. For companies that did not publish standalone sustainability reports, the sustainability-related sections of their annual reports were assessed instead. The Report_Type variable in the dataset indicates which document type was used for each observation.

Relation to Prior Research

The present dataset originates from a doctoral research program at UPI that comprises two independent measurement workstreams. The first workstream, conducted earlier and published in JDM [13], employs the Strategic ESG Capability Index (SECI) with 54 composite indicators scored on a 0–3 scale, with data collection completed in early 2024 covering FY 2023 reports for 130 firm-year observations across 18 sectors. The second workstream, described in the present article, employs the GRI 144-indicator binary disclosure framework (1/0 scoring), with data collection completed in late 2025 covering FY 2023 and FY 2024 reports for 71 forest-utilizing observations. The two workstreams use independent codebooks, training procedures, and inter-coder reliability assessments; their outputs are conceptually distinct and serve complementary purposes within the doctoral research program. The granular indicator-level data made publicly available through the present article are not contained in [13] and are not derivable from the SECI composite scores reported there.

ESG Scoring Framework

The assessment framework employs 144 indicators derived from the Global Reporting Initiative (GRI) Standards. Following the GRI Standards structure, indicators were drawn from four GRI series: Universal Standards covering General Disclosures and Management Approach (GRI 102-103, 59 items), Economic performance (GRI 201-206, 13 items), Environmental impact (GRI 301-308, 32 items), and Social responsibility (GRI 401-419, 40 items). For ESG dimension scoring, the General Disclosures and Economic indicators were combined into a Governance dimension (72 items), as economic indicators such as anti-corruption (GRI 205) and anti-competitive behavior (GRI 206) substantively reflect governance practices [1]. This yields three ESG dimensions: Governance (72 items), Environmental (32 items), and Social (40 items). The complete indicator list with GRI codes and descriptions is provided in the Indicator_List worksheet of the deposited dataset.

Each indicator was assessed using a binary scoring approach, where a score of 1 was assigned if the company disclosed information corresponding to the indicator and 0 if no relevant disclosure was identified. The assessment involved systematic content analysis through keyword searches within the source documents, following the methodology established by Phung et al. [7] and Rosman et al. [8]. Assessors searched for key terms associated with each GRI indicator (e.g., "energy consumption", "biodiversity", "anti-corruption") and evaluated whether the reported information substantively addressed the indicator requirements. Table 1 presents the detailed indicator framework with examples of assessment criteria.

Table 1. GRI indicator framework by ESG dimension.

ESG Dimension	GRI Series	No. of Indicators	Example Indicators
Governance	GRI 102: General Disclosures	56	Organizational profile, strategy, ethics, governance structure, stakeholder engagement
	GRI 103: Management Approach	3	Explanation of material topics, management approach, evaluation
	GRI 201-206: Economic	13	Economic performance (201), anti-corruption (205), anti-competitive behavior (206)
	Subtotal	72	
Environmental	GRI 301-308	32	Materials (301), energy (302), water (303), biodiversity (304), emissions (305), waste (306)
Social	GRI 401-419	40	Employment (401), OHS (403), training (404), diversity (405), human rights (406-412), community (413)
Total		144	

Dimension scores were calculated as the percentage of disclosed indicators within each dimension: $G_Score = (\text{number of Governance indicators disclosed} / 72) \times 100$; $E_Score = (\text{number of Environmental}$

indicators disclosed / 32) x 100; $S_Score = (\text{number of Social indicators disclosed} / 40) \times 100$. The composite ESG score was computed as the unweighted average of the three dimension scores: $ESG = (G_Score + E_Score + S_Score) / 3$. Equal weighting was applied following established practice in ESG disclosure research [7,8,11], as no theoretical consensus exists on the relative importance of ESG dimensions. An alternative composite measure, the total disclosure rate ($\text{Total_Disclosed} / 144 \times 100$), is also provided in the dataset for researchers preferring a non-weighted approach. Fig. 1 illustrates the scoring framework.

Quality Assurance

Several procedures were implemented to ensure data quality and reliability. First, the content analysis was conducted independently by two assessors. Each assessor reviewed the source documents and assigned binary scores for all 144 indicators for each company-year observation. Inter-coder reliability was evaluated using Cohen's kappa coefficient, yielding values ranging from 0.85 to 0.91 across the three ESG dimensions, indicating substantial to almost perfect agreement [12]. Discrepancies between assessors were resolved through joint review of the relevant report sections and discussion until consensus was reached.

Second, a systematic data verification protocol was applied following the completion of all assessments. This protocol included: (a) verification that all dimension scores were correctly calculated from the raw binary data in the GRI_Checklist; (b) confirmation that composite ESG scores matched the formula $(G + E + S) / 3$ for every observation; and (c) identification and documentation of outliers. The verification confirmed zero formula errors across all 71 observations, compared to three formula inconsistencies identified in the reference dataset by Phung et al. [7].

Third, each observation was assigned a Confidence rating (HIGH, MEDIUM, or LOW) based on the quality and accessibility of the source documents. Observations assessed from dedicated sustainability reports prepared in accordance with GRI Standards received a HIGH confidence rating, while those assessed from annual report sustainability sections received MEDIUM, and those relying on limited disclosure in annual reports received LOW. Fig. 2 presents the data collection and quality assurance workflow, and Fig. 3 displays the distribution of ESG scores across sectors.

RESULTS AND DISCUSSION

Dataset Structure

The dataset is publicly available through Mendeley Data (DOI: 10.17632/hnsb3h9353.2) and comprises two Excel files. The main file (ESG_Dataset_Final.xlsx) contains 71 company-year observations from 41 forest-utilizing companies across three sectors for fiscal years 2023 and 2024. The extended file (ESG_Dataset_Full_Repository.xlsx) provides 108 observations from 63 companies across seven sectors, enabling broader comparative analyses. Both files share an identical structure consisting of six worksheets. The first worksheet (README) provides dataset documentation, variable definitions, and citation information. The second worksheet (GRI_Checklist) contains the raw binary scores for all 144 GRI indicators for each company-year observation, organized by GRI category with rows representing indicators and columns representing individual companies. The third worksheet (ESG_Scores) provides a consolidated panel dataset with company identifiers, sector classification, ESG dimension scores, composite scores, and confidence ratings. The fourth worksheet (Company_List) presents the complete list of companies with sector classification, yearly coverage, and panel type (balanced or unbalanced). The fifth worksheet (Descriptive_Stats) contains pre-calculated summary statistics including overall descriptive statistics, sector-level comparisons, and temporal analysis for balanced-panel companies. The sixth worksheet (Indicator_List) catalogs all 144 GRI indicators with their codes, categories, ESG dimensions, and disclosure rates across the sample. This structure allows researchers to access data at multiple levels of granularity, from individual indicator responses to aggregated dimension and sector-level statistics. A complete list of primary data sources, including document types and access URLs for all 71 company-year observations, is provided in Supplementary Table S1.

Variable Definitions

The dataset contains variables organized into four categories. Company identifiers include Stock_Code (IDX ticker symbol), Company_Name, Year (fiscal year), Sector, Report_Type (indicating whether the assessment was based on a sustainability report, annual report, or both), and Confidence (assessment reliability rating based on source document quality). Raw disclosure counts include General_Disclosed, Economic_Disclosed, Environmental_Disclosed, and Social_Disclosed, reporting the number of indicators disclosed within each GRI

category. ESG dimension scores include G_Score (Governance, calculated as the proportion of disclosed items among 72 indicators), E_Score (Environmental, 32 indicators), and S_Score (Social, 40 indicators), each expressed as a percentage. The composite ESG_Composite is calculated as the unweighted average of the three dimension scores: $ESG = (G + E + S) / 3$. Additionally, Total_Disclosed reports the absolute count of indicators scored as disclosed out of 144 total indicators. Table 2 presents the sample distribution by sector and year.

Table 2. Sample distribution by sector and year.

Sector	2023	2024	Total Observations	Unique Companies
Palm Oil	18	19	37	21
Pulp and Paper	11	10	21	11
Forestry and Wood Products	4	9	13	9
Total	33	38	71	41

Sample Composition and Summary Statistics

The dataset encompasses 71 company-year observations from 41 unique companies. Of these, 30 companies appear in both fiscal years (balanced panel), while 11 companies have observations for only one year. Palm oil companies constitute the largest subsector (52.1% of observations), reflecting the sector's dominance in Indonesia's forest-commodity landscape. The temporal distribution shows an increase from 33 observations in 2023 to 38 in 2024, attributable to improved report availability for the more recent fiscal year. Table 3 presents descriptive statistics for ESG disclosure variables. The mean composite ESG score is 61.31%, with Environmental disclosure (69.81%) higher than Social (59.96%) and Governance (54.17%) disclosure. Environmental disclosure rates were consistently higher than Governance and Social disclosure rates across all three sectors. Notably, indicator-level analysis reveals that GRI 308-2 (supplier environmental impact assessment) and GRI 414-2 (supplier social impact assessment) recorded 0% disclosure across all 71 observations, indicating a complete absence of supply-chain impact disclosure in this sector—a finding with direct implications for EUDR readiness, which requires full traceability of forest-commodity supply chains.

Table 3. Descriptive statistics of ESG disclosure variables (N = 71).

Variable	Mean	Median	SD	Min	Max
ESG Composite (%)	61.31	62.45	10.02	29.56	77.01
Governance (%)	54.17	55.56	7.49	30.56	66.67
Environmental (%)	69.81	68.75	11.70	40.62	87.50
Social (%)	59.96	60.00	12.95	17.50	82.50

Table 4 presents ESG scores disaggregated by sector. Palm oil companies exhibit the highest mean ESG disclosure (63.42%), followed by forestry and wood products (60.96%) and pulp and paper (57.82%). The standard deviations range from 8.97 to 10.46, indicating moderate variation within each sector.

Table 4. ESG disclosure scores by sector.

Sector	N	Mean ESG (%)	SD	Mean E (%)	Mean S (%)	Mean G (%)
Palm Oil	37	63.42	10.12	72.47	62.91	54.88
Pulp and Paper	21	57.82	8.97	65.92	55.36	52.18
Forestry and Wood Products	13	60.96	10.46	68.51	59.04	55.34

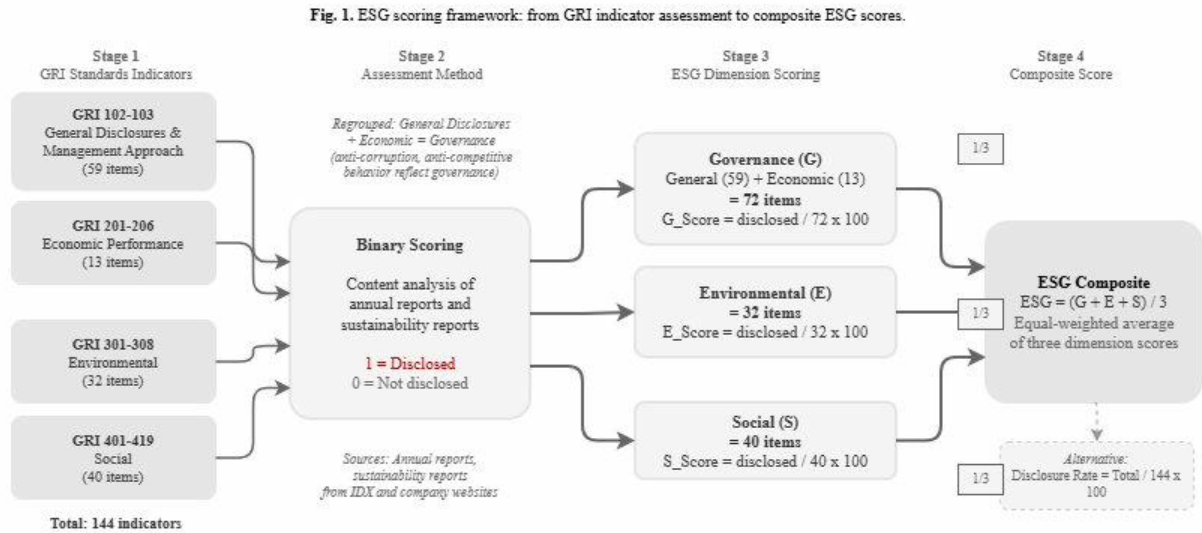


Fig. 1. ESG scoring framework: from GRI indicator assessment to composite ESG scores.

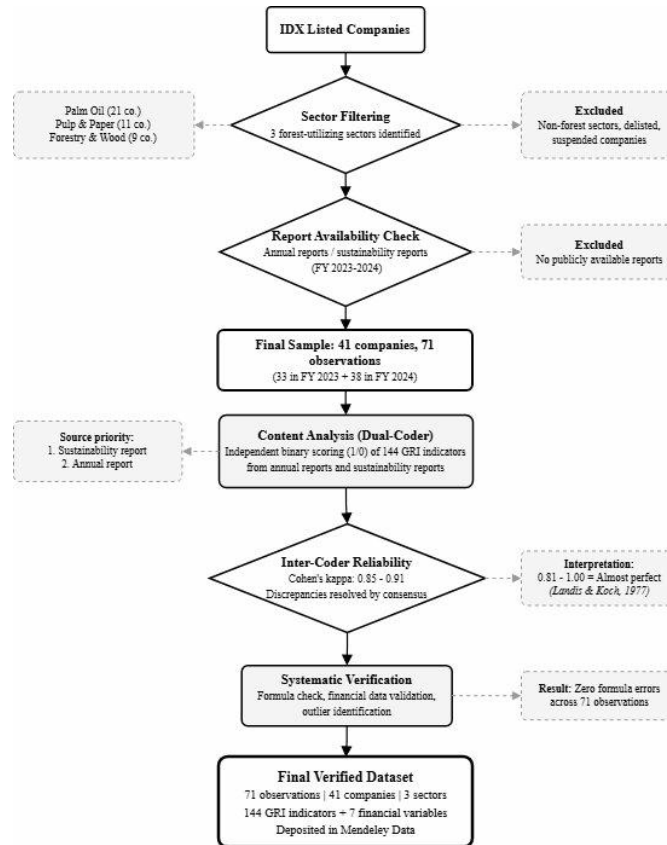


Fig. 2. Data collection and quality assurance workflow.

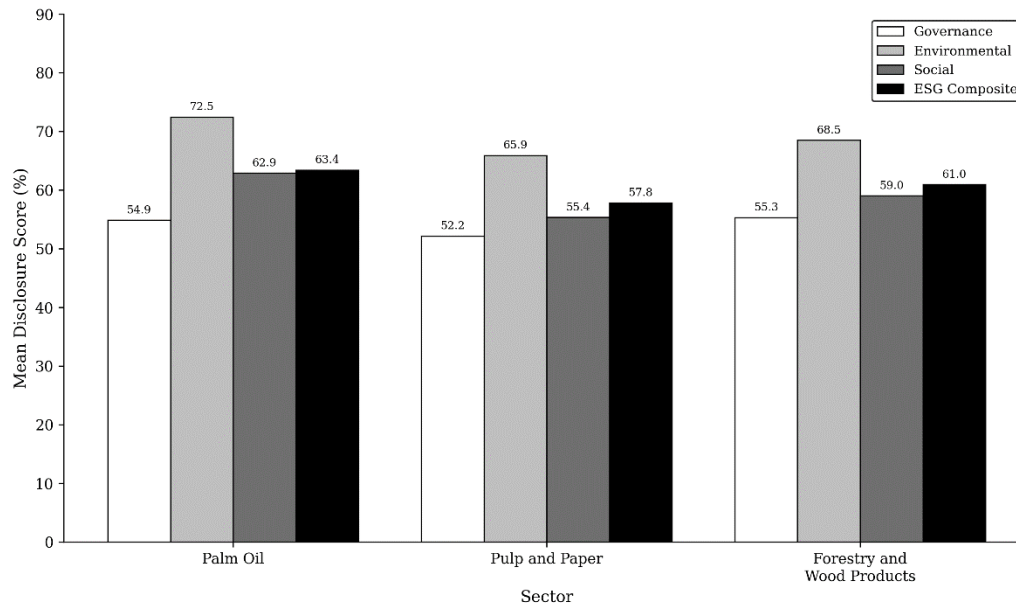


Fig. 3. Mean ESG disclosure scores (%) by sector and dimension. Bars represent Governance (white), Environmental (light grey), Social (dark grey), and ESG Composite (black). Values shown above each bar.

Limitations

Several limitations should be acknowledged when using this dataset. First, the binary scoring approach (1/0) captures the presence or absence of disclosure but does not assess the quality, depth, or verifiability of the reported information. Future research could employ weighted or multi-level scoring schemes to capture disclosure quality. Second, the dataset is limited to IDX-listed companies; privately held companies and smallholder operations, which constitute a substantial portion of Indonesia’s forest-commodity sector, are not included. Third, ESG assessments rely on corporate self-reporting, which may be subject to impression management or selective disclosure. Fourth, the Governance dimension combines GRI Universal Standards with Economic indicators, which may not align with governance definitions used in other ESG frameworks. Fifth, the dataset focuses exclusively on disclosure metrics and does not include financial performance variables; researchers interested in ESG-financial performance relationships will need to supplement these data with financial filings.

CONCLUSION

This article presents an original GRI-based ESG disclosure dataset for Indonesian forest-utilizing companies, comprising 71 company-year observations from 41 unique companies across the palm oil, pulp and paper, and forestry and wood products sectors for fiscal years 2023 and 2024. The dataset was constructed through systematic dual-coder content analysis using 144 GRI indicators, with inter-coder reliability yielding Cohen’s kappa values of 0.85 to 0.91. Descriptive findings indicate a hierarchical disclosure pattern (Environmental > Social > Governance) consistent across all three sectors, with mean composite ESG disclosure of 61.31%. Indicator-level analysis reveals 0% disclosure rates for GRI 308-2 and GRI 414-2 (supplier environmental and social impact assessment) across the entire sample, identifying a critical gap in supply-chain transparency disclosure with direct implications for EUDR readiness. The dataset is publicly deposited at Mendeley Data (DOI: 10.17632/hnsb3h9353.2) under a CC BY 4.0 license, enabling reuse for analyses of disclosure determinants, subsector comparisons, sustainability certification effectiveness, and benchmarking corporate transparency against regulatory requirements. Future research can build on this dataset to examine the relationship between ESG disclosure and corporate financial performance, assess the impact of EUDR enforcement on disclosure trajectories, and develop weighted or quality-adjusted disclosure scoring schemes.

ACKNOWLEDGMENTS

This research was conducted as part of a doctoral research program at Universitas Pendidikan Indonesia (UPI). The authors gratefully acknowledge the Indonesia Stock Exchange (IDX) for providing public access to corporate filings and reports. The dataset described in this article is publicly deposited at Mendeley Data (DOI: 10.17632/hnsb3h9353.2, CC BY 4.0).

REFERENCES

- [1] Global Reporting Initiative. GRI Standards. Amsterdam: GRI; 2021. Available from: <https://www.globalreporting.org/standards/>
- [2] Otoritas Jasa Keuangan (OJK). Peraturan OJK Nomor 51/POJK.03/2017 tentang Penerapan Keuangan Berkelanjutan bagi Lembaga Jasa Keuangan, Emiten, dan Perusahaan Publik. Jakarta: OJK; 2017.
- [3] Benedict J, Biddle H, Gollnow F, Heilmayr R, Mueller C, Ribeiro V. Indonesia palm oil (2021-2022) (Version 1.2.2) [Data set]. Trase/Stockholm Environment Institute; 2024. <https://doi.org/10.48650/X83N-7M36>
- [4] Directorate General of Plantations, Ministry of Agriculture. Indonesian Palm Oil Statistics 2023. Jakarta: Ministry of Agriculture; 2024.
- [5] Regulation (EU) 2023/1115 of the European Parliament and of the Council of 31 May 2023 on the making available on the Union market and the export from the Union of certain commodities and products associated with deforestation and forest degradation. Official Journal of the European Union. 2023;L150:206-247.
- [6] Zoological Society of London (ZSL). SPOTT Palm Oil Assessment Summary -- November 2024. London: ZSL; 2024. Available from: <https://www.spott.org/palm-oil-assessment-summary/>
- [7] Phung MT, Dao VT, Mai KT. Dataset for analysing the ESG-oriented technical efficiency of VNSI listed companies. Data in Brief. 2024;52:109832. <https://doi.org/10.1016/j.dib.2023.109832>
- [8] Rosman NS, Ho WK, Hashim HA, Devi KS, Kanagasabapathy S, Singh J. Environmental and Social Disclosures Dataset for Malaysian Public Listed Companies. Data in Brief. 2023;50:109463. <https://doi.org/10.1016/j.dib.2023.109463>
- [9] Gutierrez-Ponce H, Wibowo SA. Sustainability Reports and Disclosure of the Sustainable Development Goals (SDGs): Evidence from Indonesian Listed Companies. Sustainability. 2023;15(24):16919. <https://doi.org/10.3390/su152416919>
- [10] Soleha N, Rosiana R. Exploring Sustainability Reports in the Indonesian Palm Oil Industry. Journal of Applied Business, Taxation and Economics Research. 2021;1(1):77-94. <https://doi.org/10.54408/jabter.v1i1.23>
- [11] Hanapi NH, Alahdal WM, Hashim HA, Salleh Z, Taha R, Adnan SM. Datasets for Sustainability Reporting Index of Agricultural and Plantation Companies in Malaysia. Data in Brief. 2025;60:111491. <https://doi.org/10.1016/j.dib.2025.111491>
- [12] Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159-174. <https://doi.org/10.2307/2529310>
- [13] Pragama TD, Disman, Hendrayati H, Dirgantari PD. From Ambidexterity to Organizational Agility: The Mediating Role of Strategic Integration in ESG Context. JDM (Jurnal Dinamika Manajemen). In press (accepted for Vol. 17(1), 2026). Available from: <https://journal.unnes.ac.id/journals/jdm>