Journal of Auditing, Finance, and Forensic Accounting

Volume 14
Issue 1

Available Online at: https://journal.trunojoyo.ac.id/jaffa

Comparing Beneish and Dechow Model in Predicting Financial Shenanigans: Evidence from Energy Sector in Emerging Market

Defel Septian, Alfian Sayuti, Khairunnisa, Ayu Ambang Lestari, *Vina Febrianti
Universitas Bumigora, Indonesia

Article Info:
Received: *February, 2026*
Revised: *April, 2026*
Accepted: *April, 2026*

DOI:
10.21107/jaffa.v14i1.33632

ISSN (Print):
2339-2886

ISSN (Online):
2461-0607

Page:
17-33

ABSTRACT
Based on a 2024 ACFE survey, financial statement fraud is the type of fraud that causes the highest average loss compared to other types. Several approaches exist for predicting financial statement fraud, including the Beneish M-Score (hereafter M-Score) and Dechow F-Score (hereafter F-Score). The purpose of this study is to evaluate which model best predicts financial statement fraud in Indonesia's energy industry. The research sample was determined using a purposive sampling method. The sample consists of 140 financial statements from energy sector companies listed on the Indonesia Stock Exchange from 2020 to 2024 that were indicated to be free from violations related to their financial reporting. The results reveal that the F-Score model is more effective in detecting financial statement fraud in energy companies, with a 100% accuracy rate and a 0% error rate, while the M-Score model obtained a 73% accuracy rate and a 27% error rate. This research is expected to provide implications for stakeholders in selecting the right model to predict financial statement fraud, especially in energy sector.

Keywords:
Beneish M-Score, Energy Sector, Dechow F-Score, Financial Statement Fraud, Prediction Model.

INTRODUCTION

Fraud, particularly at the financial statement level, represents a critical challenge within the global business landscape (Azis et al., 2025; Haldar et al., 2025; Halteh & Tiwari, 2023; MacCarthy, 2017; Madah Marzuki et al., 2024), garnering extensive scrutiny from scholars, practitioners, and relevant stakeholders (Bansal, 2024; Ghorbel & Ben Amar, 2025; Kukreja et al., 2020). Within the context of Indonesia as an emerging economy, financial statement manipulation is a pervasive phenomenon (Aviantara, 2023). This specific category of misconduct is regarded as the most destructive, yielding profound financial repercussions (Biduri & Tjahjadi, 2026; Himmah et al., 2025; Khamainy et al., 2022; Shahzadi et al., 2025).

Research into the detection of financial statement fraud is paramount, as it mitigates the information costs incurred by decision-makers during the investment process (Beneish & Vorst,

Copyright © 2026 Authors. This is an open access article under the CC-BY license.

2022). Such illicit practices yield profound systemic consequences, ranging from direct financial losses and diminished productivity to heightened job insecurity and the erosion of public confidence in capital markets (Demetriades & Owusu-Agyei, 2022; du Toit, 2024; Sylwestrzak & Słomkowska, 2025). Consequently, an extensive body of literature has evaluated the efficacy of diverse detection methodologies. For instance, Aboud & Robinson (2022) contrasted traditional financial ratio analysis with advanced Artificial Intelligence (AI) and data mining techniques. Similarly, Aviantara (2023) scrutinized various bankruptcy and fraud prediction frameworks within the context of Garuda Indonesia. Furthermore, Beneish & Vorst (2022) provided a comprehensive comparative analysis of seven distinct models including the M-Score, Kernel Model, F-Score, Adjusted Benford Score, Alawadhi's Misrepresentation Model, and Bao's Machine Learning-based predictions. Other scholars have specifically focused on the comparative performance of the M-Score and the F-Score in identifying reporting irregularities (Aghghaleh et al., 2016; Kaab Omeir et al., 2023; Ratmono et al., 2020; Sylwestrzak & Słomkowska, 2025).

Doss & Bacha (2026) contend that conventional financial ratio-based detection frameworks are increasingly obsolete, as they often fail to incorporate broader institutional nuances and multifaceted accounting dimensions. Conversely, while contemporary techniques leveraging AI and data mining offer advanced capabilities, they are frequently constrained by their exclusivity, high implementation costs, and operational complexity (Aboud & Robinson, 2022). Consequently, ratio-based models most notably the M-Score and F-Score retain their prevalence in academic and professional practice for several reasons: (1) their reliance on transparent, readily available financial disclosures provides high interpretability and significant diagnostic value (Biduri & Tjahjadi, 2026; Javaid et al., 2025); (2) they offer a cost-effective and parsimonious alternative to resource-intensive methods (du Toit, 2024; Maniatis, 2022; Repousis, 2016); and (3) empirical evidence suggests that both models yield tangible net benefits for investors in the proactive identification of reporting irregularities (Beneish & Vorst, 2022).

The landscape of corporate malfeasance in Indonesia is characterized by escalating complexity, particularly within strategic sectors such as energy. A prominent illustration is the high-profile corruption case involving PT Pertamina, the nation's primary oil

Table 1. Fraud Types and Average Loss

Fraud Types	Average Loss	Case Percentage
Asset Misappropriation	$120,000	89%
Corruption	$200,000	48%
Financial Statement Fraud	$766,000	5%

Source: ACFE, 2024

and energy entity, which centered on procurement irregularities and resulted in estimated state losses of approximately IDR 285 trillion (Attorney General's Office, 2025). Such incidents underscore the urgent mandate for corporate governance to develop robust mechanisms for the early detection and mitigation of fraudulent financial reporting. Similar occurrences of financial manipulation have been documented at PT Timah Tbk (TINS), further highlighting systemic vulnerabilities in the industry (Antoni, 2025; Sary & Rani, 2025). In this context, financial statement analysis serves as a critical diagnostic tool, providing a systematic framework to identify high-risk areas and red flags indicative of potential fraud (Singleton & Singleton, 2010:32). Predominant among the methodologies employed to assess the probability of financial statement fraud are the M-Score and the F-Score models, both of which are established as standard empirical proxies in the contemporary fraud literature (Himmah et al., 2025; Kaab Omeir et al., 2023; Ratmono et al., 2020; Soepriyanto et al., 2021; Sylwestrzak & Słomkowska, 2025).

Extant literature has frequently employed these models concurrently to evaluate their relative diagnostic efficacy. For instance, Zhou et al. (2022) demonstrated that the M-Score model yielded statistically consistent results in identifying fraudulent financial reporting, whereas the F-Score model exhibited insufficient predictive power and was deemed less suitable for fraud detection. Similarly, Masitah (2024) reported that the M-Score achieved a superior accuracy rate of 84.6% compared to 82.1% for the F-Score, suggesting that the former is a more effective predictor of financial statement fraud. However, the cross-border applicability of these frameworks remains a subject of debate; Ye et al. (2022) observed that neither model attained absolute precision, primarily attributing these discrepancies to the divergence in accounting standards between the United States and China. Despite these limitations, both models remain instrumental as preliminary diagnostic instruments for detecting potential financial statement fraud.

Stream of research reveals significant disparities regarding the superior methodology for detecting financial statement fraud, a phenomenon that Aboud & Robinson (2022) attribute to historical and geographical heterogeneities. Furthermore, du Toit (2024) contends that the idiosyncratic characteristics of fraud vary across cases, while industry-specific nuances further exacerbate the complexity of identifying such unethical practices. Prior scholarly inquiries have predominantly focused on multi-country analyses (Ghorbel & Ben Amar, 2025; Hasan et al., 2017), cross-sectoral evaluations (Doss & Bacha, 2026; Madah Marzuki et al., 2024; Maniatis, 2022; Repousis, 2016; Shahzadi et al., 2025; Soepriyanto et al., 2021), or specific industries such as aviation (Aviantara, 2023), finance (Biduri & Tjahjadi, 2026; Himmah et al., 2025; Khatun et al., 2022), technology (Demetriades & Owusu-Agyei, 2022), and manufacturing (Khamainy et al., 2022). Consequently, the present study addresses this critical empirical lacuna by

examining the Indonesian context, specifically focusing on the energy sector an area that remains under-researched despite its susceptibility to financial statement fraud.

Based on an empirical analysis of 140 financial observations from 91 Indonesian energy sector firms spanning the period 2020–2024 within a sample exhibiting no documented reporting violations this study demonstrates that the F-Score serves as a superior methodology for detecting potential financial statement fraud in this specific context. These results address a notable empirical void and advance the extant literature on fraud detection in several ways. Theoretically, this research provides evidence that the F-Score outperforms the M-Score in identifying financial statement fraud within the energy sector of an emerging economy, namely Indonesia. Practically, these findings suggest that investors, corporate entities, and regulatory bodies in Indonesia should leverage the F-Score as a primary diagnostic tool. Furthermore, this study advocates for the institutionalization of such fraud-detection analyses, proposing their inclusion as mandatory supplementary disclosures within annual financial reports to enhance market transparency.

LITERATURE REVIEW AND HYPHOTESIS DEVELOPMENT

Financial Statement Fraud

Consistent with the frameworks established by ACFE (2024) and Maniatis (2022), financial statement fraud is conceptualized as a deliberate deception or misrepresentation orchestrated by individuals or entities, predicated on the knowledge that such distortions will yield unauthorized benefits to the organization or external stakeholders. Substantively, this phenomenon involves the manipulation of financial disclosures to artificially inflate profitability beyond actual performance levels. Such malfeasance is typically executed by senior management, driven by a complex interplay of incentives including personal enrichment, the imperative to ensure organizational solvency, and the pressure to maintain professional prestige and leadership status (Abusharbeh & Zakarneh, 2024; Biduri & Tjahjadi, 2026; Demetriades & Owusu-Agyei, 2022; Hasan et al., 2017; Khatun et al., 2022).

Tarjo & Herawati (2021) conceptualize financial statement fraud through the 'three Ms' framework: the manipulation or alteration of accounting records, the omission of material transactions or events, and the misapplication of accounting principles regarding classification and disclosure. Amidst the rising prevalence of such fraudulent activities over recent decades (Kukreja et al., 2020), there has been a corresponding shift toward more advanced detection methodologies and preventative frameworks (Aboud & Robinson, 2022; Kaab Omeir et al., 2023; Ramzan & Lokanan, 2025).

The prevalence of financial statement fraud in Indonesia spans multiple industries (Khamainy et al., 2022), yet detection is often hindered by restricted investigative resources and the adept use of technology to exploit regulatory gaps (Biduri & Tjahjadi, 2026;

Cahyono et al., 2025). High-profile cases over the last five years, such as those involving PT Timah Tbk (TINS) and PT Perusahaan Gas Negara Tbk (PGAS), exemplify these risks (Antoni, 2025; Sary & Rani, 2025). Evidence from the TINS scandal, which incurred IDR 271 trillion in state losses, reflects profound weaknesses in corporate integrity (Antoni, 2025; Sary & Rani, 2025). Similarly, allegations surrounding PGAS reveal critical fraud exposure within the operational and reporting frameworks of Indonesian state-owned entities.

Beneish M-Score Model

The Beneish M-Score is a robust analytical framework designed to identify potential financial statement manipulation. Originally developed by Beneish (1999) in the seminal study 'The Detection of Earnings Manipulation,' the model differentiates between manipulators and non-manipulators by analyzing quantitative variances across specific financial ratios. This historical data-driven approach serves as a critical deductive instrument not only for identifying earnings management but also for predicting broader financial irregularities, such as money laundering (Cahyono et al., 2025). Consequently, the M-Score remains a vital tool for enhancing financial reporting quality and preempting future fraudulent activities (Abusharbeh & Zakarneh, 2024).

This model utilizes a composite of financial ratios to generate a specific score, facilitating the identification of fraudulent indicators with global applicability (Khatun et al., 2022). Typically, entities engaging in earnings manipulation exhibit characteristic anomalies, such as disproportionate revenue surges coupled with significant declines in expense accounts within a single reporting period. The M-Score framework operationalizes these observations through eight distinct indicators focusing on asset fluctuations and revenue trajectory (Hasan et al., 2017). These include the Days Sales in Receivable Index (DSRI), Gross Margin Index (GMI), Asset Quality Index (AQI), Sales Growth Index (SGI), Depreciation Index (DEPI), Sales General and Administrative Expenses Index (SGAI), Leverage Index (LVGI), and Total Accruals to Total Assets (TATA).

Dechow F-Score Model

The F-Score model serves as an evolution of the M-Score, distinguished by its inclusion of SEC AAERs from 1982 through 2005, whereas the original M-Score was limited to releases ending in 1992 (Dechow et al., 2011). This temporal expansion provides the F-Score with superior predictive relevance due to its reliance on more exhaustive and recent regulatory data (Azis et al., 2025). Developed through scaled logistic probability modeling, the F-Score employs a rigorous variable selection process to enhance its diagnostic capabilities in identifying financial statement fraud (Dechow et al., 2011).

Dechow et al. (2011) developed three logistic regression models to assess fraud risk based on accrual quality, financial performance, and market-based measures. Despite the inclusion of non-financial and off-balance sheet variables in Models II and III,

Model I remains the most statistically robust version for predicting financial statement fraud (Dechow et al., 2011). Adopting this parsimonious approach, the present study utilizes Model I, focusing on seven fundamental financial ratios: RSST Accruals, change in receivables, change in inventories, soft assets, change in cash sales, change in return on assets, and issuance activities (Dechow et al., 2011).

RESEARCH METHODS

This study employs a comparative quantitative research design, a methodology structured to evaluate similarities and discrepancies across multiple groups, conditions, or variables (Suganda, 2025). This approach is particularly effective in determining whether specific conditions yield divergent outcomes across distinct cohorts. The research population comprises energy sector companies listed on the Indonesia Stock Exchange (IDX) for the 2020–2024 period. The selection of this sector is strategically justified by three factors: (1) the relative scarcity of sector-specific fraud literature; (2) the high-profile nature of recent financial irregularities within this industry (Antoni, 2025; Sary & Rani, 2025); and (3) the escalating environmental impact of its operations, which has shifted investor scrutiny toward the intersection of financial integrity and sustainability initiatives (Ghorbel & Ben Amar, 2025).

This research employs a purposive sampling technique, drawing from an initial population of 91 entities. From this, a final sample of 140 financial reports was selected based on the specific criterion of energy sector companies with no recorded reporting violations. Both the M-Score and F-Score predictive frameworks were utilized as diagnostic instruments to detect potential financial statement fraud. These models serve as the foundational metrics for a comparative analysis intended to ascertain their relative efficacy in fraud detection. The operationalization of the M-Score necessitates the preliminary computation of eight distinct financial ratios, which are subsequently integrated into the primary model as follows:

a. Days Sales in Receivable Index (DSRI)
   DSRI= ((Account Receivables (t))/ Sales (t))/ ((Account Receivables (t-1))/ Sales (t-1))

b. Gross Margin Index (GMI)
   GMI= ((Sales(t-1)-Cost of Goods Sold(t-1))/ Sales(t-1))/ ((Sales(t)-Cost of Goods Sold(t))/ Sales(t))

c. Asset Quality Index (AQI)
   AQI= ((1-Current Assets(t)+Net Fixed Assets(t))/(Total Assets(t))) / ((1-Current Assets(t-1)+Net Fixed Assets(t-1))/(Total Assets(t-1)))

d. Sales Growth Index (SGI)
   SGI=(Sales(t))/ (Sales(t-1))

e. Depreciation Index (DEPI)
DEPI=((Depreciation(t-1))/(Net Fixed Assets(t-1)+Depreciation(t-1)))/((Depreciation(t))/(Net Fixed Assets(t)+Depreciation(t)))

f. Sales, General and Administrative Expenses Index (SGAI)
SGAI=((SGA(t))/(Sales(t)))/((SGA(t-1))/(Sales(t-1)))

g. Leverage Index (LVGI)
LVGI=(Current Liabilities(t)+Total Long-Term Debt(t)+Total Assets(t))/(Current Liabilities(t-1)+Total Long-Term Debt(t-1)+Total Assets(t-1))

h. Total Accrual to Total Assets (TATA)
TATA=(Net Operating Profit(t)-Cash Flow from Operating(t))/(Total Assets(t))

According to Beneish (1999), there are eight index ratios that make up the M-Score. The calculation results are then formulated using the following measurement formula:

M-Score = -4,840+0,920(DSRI)+0,528(GMI)+0,404(AQI)+0,892(SGI) + 0,115(DEPI) – 0,172(SGAI) – 0,327(LVGI) + 4,697(TATA)

Classification within the M-Score model is based on a -2.22 cut-off value, where scores greater than this threshold identify potential manipulators and lower scores denote non-manipulators (Beneish, 1999; Hasan et al., 2017; Khamainy et al., 2022). Regarding the F-Score methodology, the diagnostic process involves a multi-stage calculation of seven independent variables (Dechow et al., 2011). These variables, which must be computed before the final model application, include:

a. RSST Accrual (Richardoson, Sloan, Soliman & Tuna)
RSSTA=(ΔWC+ΔNCO+ΔFIN)/(Average Total Assets)

b. Change in Receivable (REC)
REC=ΔReceivables/(Average Total Assets)

c. Change in Inventory (INV)
INV=ΔInventory/(Average Total Assets)

d. Soft Assets (SOFTASSETS)
SOFTASSETS=(Total Assets-Net Fixed Assets-Cash and Cash Equivalent)/(Total Assets)

e. Change in Cash Sales (CASHSALES)
CASHSALES=(Sales(t)-ΔReceivables(t))/(Sales(t-1)-ΔReceivables(t-1))

f. Return on Assets (ROA)
ROA=((Earnings(t))/(Average Total Assets(t)))/((Earnings(t-1))/(Average Total Assets(t-1)))

g. Actual Issuance of Stock (ISSUE)
Issurance is worth 0 or 1 (1 if the company issued bonds or common stock/preferred stock in the period in question, and 0 otherwise).

After calculating all F-Score variables, the results are entered into the predicted value equation or predicted value determined by Dechow et al. (2011).

Predicted Value = -7,893 + 0,790 (RSST) + 2,518 (ΔREC) + 1,191 (ΔINV) + 1,979 (SOFTASSETS) + 0,171 (ΔCASHSALES) – 0,932 (ΔROA) + 1,029 (ISSUE)

Predicted value is converted into probability value, with the formula:

Probability Value = e^(Predicted Value) / (1+e^(Predicted Value)

The F-score is derived by dividing the predicted probability by an unconditional misstatement probability of 0.0037. Within this framework, a threshold of 1.00 is established; an F-score exceeding this benchmark serves as a diagnostic indicator of potential financial statement fraud (Kaab Omeir et al., 2023). Furthermore, the probability of fraudulent reporting is positively correlated with the magnitude of the F-score, signifying that higher values denote an increased risk of manipulation (Kaab Omeir et al., 2023).
F-Score = Probability Value / Unconditional Probability

Prediction Model Accuracy Level
Upon completing the calculations for the entire sample, the predicted instances of both fraudulent and non-fraudulent financial reporting will be aggregated. These results facilitate the determination of each model's predictive accuracy in identifying financial statement fraud (Mulya et al., 2024). Following the methodological frameworks established by Aghghaleh et al. (2016) and Kaab Omeir et al. (2023), the accuracy of the respective models is quantified using the following formula:

Accuracy Level= ((Number of Non-Fraud / Non-Manipulator Predictions) / Number of Samples) × 100%

The frequency of non-fraud predictions represents the cohort of energy companies officially cleared of financial irregularities by the Indonesia Stock Exchange, where the M-Score and F-Score results align with these official designations (Islamiyati & Purnomo, 2021). Conversely, the error rate quantifies the incidence of misclassification when the models flag a company for financial statement fraud despite a clean regulatory record. Specifically, the volume of fraud predictions identifies samples officially declared non-fraudulent by the exchange but classified as fraudulent through the M-Score and F-Score analytical frameworks (Mulya et al., 2024). This error rate, or Type I error, is determined using the following formula:

Error Level= ((Number of Fraud/ Manipulator Predictions)/ Number of Samples)/ × 100%

The volume of fraud predictions corresponds to instances where energy companies, officially cleared of financial irregularities by the Indonesia Stock Exchange, are nonetheless flagged as fraudulent by the M-Score and F-Score frameworks. Models exhibiting robust accuracy coupled with diminished error rates are deemed more efficacious for fraud detection, as they demonstrate a superior capacity to differentiate between manipulative and non-manipulative reporting practices (Islamiyati & Purnomo, 2021).

RESULTS AND DISCUSSION

Financial statement fraud remains a paramount concern due to its direct association with managerial misconduct and its deleterious impact on investor interests (du Toit, 2024; Madah Marzuki et al., 2024; Ramzan & Lokanan, 2025; Sylwestrzak & Słomkowska, 2025). Consequently, robust diagnostic frameworks such as the M-Score and F-Score models have been developed to enhance detection capabilities (Abusharbeh & Zakarneh, 2024; Azis et al., 2025; Ghorbel & Ben Amar, 2025; Himmah et al., 2025). Within these frameworks, an M-Score exceeding the -2.22 threshold (Hasan et al., 2017; Khamainy et al., 2022) or an F-Score surpassing 1.00 (Kaab Omeir et al., 2023) serves as an indicator of potential manipulation. Crucially, the probability of fraudulent reporting is positively correlated with the magnitude of these scores; as values escalate, the likelihood of material misstatement increases. Both models have demonstrated high empirical accuracy in forecasting the prevalence of financial irregularities (Aghghaleh et al., 2016; Ratmono et al., 2020).

Table 2 delineates the comparative accuracy of the predictive frameworks within the Indonesian energy sector. The empirical findings reveal that the F-Score model achieved an accuracy rate of 100% with a negligible error rate (0%), indicating total alignment with the non-fraudulent classifications established by the IDX. In contrast, the M-Score model yielded a markedly lower accuracy rate of 73% and an associated error rate of 27%, underscoring its relative diagnostic limitations in this specific sectoral context. Consequently, these results suggest that the F-Score model possesses superior predictive efficacy and provides a more reliable mechanism for detecting financial statement fraud than the M-Score framework.

Table 2. Accuracy Level of M-Score and F-score Models

Prediction	M-Score	F-Score
Non-Manipulator/Non-Fraudulent	102	140
Manipulator/ Fraudulent	38	0
Total Sample	140	140
Accuracy Level	73%	100%
Error Level	27%	0%

Source: Data Processed

This study concludes that the F-Score model exhibits superior efficacy compared to the M-Score framework, particularly regarding its capacity to generate consistent, stable outcomes that comprehensively capture fraud risk. While the M-Score maintains utility as a preliminary screening tool, the F-Score emerges as the more robust primary instrument for analyzing structural and persistent irregularities. Notably, these findings diverge from the empirical evidence presented by Husnurrosyidah & Fatihah (2022), who identified the M-Score as the more effective predictive model. Furthermore, this research contradicts the conclusions of Abusharbeh & Zakarneh (2024), Aviantara (2023), Beneish & Vorst (2022), Kaab Omeir et al. (2023), and Ye et al. (2022), all of whom argued for the M-Score's superior performance in detecting financial statement fraud.

The findings of this study resonate with the evidence presented by Sitoresmi et al. (2024) and Andrew et al. (2022), both of whom identify the F-Score as a highly efficacious instrument for the early detection and prediction of financial statement fraud. The perceived superiority of the F-Score may be attributed to its multi-dimensional architecture, which integrates financial and non-financial metrics spanning accrual quality, operational performance, off-balance sheet activities, and market-based indicators to provide a comprehensive assessment of potential earnings overstatement (Sylwestrzak & Słomkowska, 2025). While inherent limitations persist regarding the predictive accuracy of any single model (Kukreja et al., 2020), the F-Score demonstrated higher diagnostic precision within the Indonesian energy landscape. This performance may stem from the observation that Indonesian energy firms often eschew accrual-based earnings management (Bansal, 2024), potentially rendering conventional detection methods less effective. Nevertheless, these preliminary findings necessitate further empirical validation in future longitudinal studies.

The efficacy of the F-Score model is demonstrated by its longitudinal stability and its capacity to structurally encapsulate the risk of material misstatement. This finding is rooted in the distinct conceptual objectives of the two frameworks; while the M-Score is designed to detect transient, short-term earnings manipulation, the F-Score serves as a medium-term diagnostic for systemic fraud risk. Its superiority in this context stems from its constituent variables notably RSST accruals, the proportion of soft assets, and indicators of financial pressure which possess a higher degree of persistence. Unlike indicators sensitive to temporary operational volatility or discretionary accounting shifts, these variables reflect fundamental corporate conditions. Consequently, the F-Score provides a more consistent risk classification across periods, reinforcing the premise that financial statement fraud is a structural phenomenon that evolves incrementally rather than an isolated, instantaneous occurrence.

Conceptually, the F-Score is predicated on a probabilistic framework, distinguishing it from the threshold-centric ratio indices characteristic of the M-Score. Within this model, the unconditional probability of misstatement establishes the baseline propensity for financial irregularities across the broader corporate population. This baseline serves as a primary reference point for risk assessment, which is iteratively refined through conditional probabilities derived from an entity's specific accounting and financial signals. Consequently, the F-Score avoids the categorical classification of every financial anomaly as fraud; instead, it evaluates whether such variances significantly elevate the probability of misstatement above the population norm. Methodologically, this probabilistic approach yields more conservative and stable outcomes by requiring a confluence of significant indicators to shift the risk profile away from the baseline. This inherent conservatism reduces the incidence of false positives particularly for entities experiencing non-systemic fluctuations thereby resulting in a lower misclassification rate compared to the M-Score.

While the M-Score emphasizes continuous-number forensic analysis for fraud detection, the F-Score is primarily oriented toward identifying accrual-based earnings management (Bansal, 2024). The integration of an unconditional probability of misstatement necessitates an inherent trade-off between diagnostic sensitivity and specificity. Specifically, the F-Score exhibits lower sensitivity but superior specificity, prioritizing the accurate identification of non-fraudulent entities (Dechow et al., 2011). This conservative classification logic ensures that fraud is only flagged when the calculated probability surpasses a rigorous statistical threshold. Consequently, while the F-Score significantly minimizes the incidence of false positives (Type I errors), it may increase the risk of false negatives (Type II errors) in cases of nascent or non-systemic fraudulent activity.

In contrast, the M-Score model in this study demonstrates heightened sensitivity, characterized by significant inter-period volatility in its output scores. This sensitivity is largely attributable to the model's reliance on ratio-based variables such as the DSRI, AQI, and SGI which are inherently procyclical and highly responsive to business cycle dynamics and corporate growth trajectories. While such responsiveness serves as a valuable early warning mechanism, the empirical evidence suggests that the M-Score is more susceptible to accounting noise, often generating signals of manipulation that lack longitudinal sustainability. Consequently, this volatility diminishes the model's efficacy as a primary instrument for fraud risk analysis, where consistency and inferential precision are paramount.

CONCLUSIONS AND SUGGESTIONS

This study evaluates the comparative suitability of two prominent fraud detection models within the specific context of the Indonesian energy sector. Utilizing a sample of 140 financial reports from energy companies officially cleared of reporting violations by the IDX between 2020 and 2024, the empirical findings reveal significant divergence in model performance. The M-Score framework classified 102 reports as non-manipulative and 38 as potentially manipulative. In contrast, the F-Score model yielded 140 non-fraudulent classifications, representing a 100% alignment with official regulatory data. Consequently, the F-Score demonstrated superior diagnostic efficacy for the IDX-listed energy sector. Although certain entities exhibited fluctuations in isolated financial ratios, the integrated analysis of RSST accruals, soft asset proportions, and financial pressure did not manifest a collective pattern indicative of systematic financial statement fraud.

This study offers significant theoretical and practical contributions to the literature on financial statement fraud detection. First, it provides robust empirical evidence demonstrating the superior diagnostic precision of the F-Score over the M-Score within the Indonesian energy sector. This finding effectively addresses a critical research gap regarding fraud detection in emerging market energy industries. Second, integrating F-Score-based analysis into corporate annual reports could substantially enhance the transparency and credibility of financial disclosures. For investors, the F-Score serves as a pragmatic tool for obtaining additional assurance regarding an entity's financial integrity. Finally, to bolster systemic confidence among capital market participants, policymakers should consider formalizing regulations that mandate the inclusion of F-Score predictive analytics within the reporting frameworks of publicly listed companies.

Despite the insights provided, the findings of this study should be interpreted with caution in light of several limitations. First, the generalizability of these results is constrained to the Indonesian energy sector, given the specific sectoral justifications employed; thus, future research should explore these models across diverse industries and geographical contexts, or through cross-sectoral comparative analyses. Second, this study did not employ auxiliary statistical validations to further substantiate the robustness of the findings. Subsequent inquiries could integrate difference-in-means tests or rigorous regression analyses to strengthen the empirical distinction between manipulative and non-manipulative entities. Finally, this investigation was limited to two ratio-based fraud detection models, selected for their cost-effectiveness and accessibility from an investor's perspective. Future studies should broaden this scope by adopting the perspectives of corporate management or regulatory bodies and conducting comparative accuracy assessments between ratio-based models, data mining techniques, and text-mining methodologies.

REFERENCES

Aboud, A., & Robinson, B. (2022). Fraudulent Financial Reporting and Data Analytics: An Explanatory Study from Ireland. *Accounting Research Journal, 35*(1), 21–36. https://doi.org/10.1108/ARJ-04-2020-0079.

Abusharbeh, M., & Zakarneh, E. N. (2024). Earnings Manipulation Behavior in the Non-Financial Listed Firms of Palestine: The Implication of Beneish Model. *EuroMed Journal of Business.* https://doi.org/10.1108/EMJB-05-2024-0121.

ACFE. (2024). *Report to the Nations: on Occupational Fraud and Abuse.* Association of Certified Fraud Examiners (ACFE). https://acfe.com (Accessed 17 December 2024).

Aghghaleh, S. F., Mohamed, Z. M., & Rahmat, M. M. (2016). Detecting Financial Statement Frauds in Malaysia: Comparing the Abilities of Beneish and Dechow Models. *Asian Journal of Accounting and Governance, 7,* 57-65. http://dx.doi.org/10.17576/AJAG-2016-07-05.

Andrew, A., Candy, C., & Robin, R. (2022). Detecting Fraudulent of Financial Statements Using Fraud S.C.O.R.E Model and Financial Distress. *International Journal of Economics, Business and Accounting Research (IJEBAR), 6*(1), 696–707. https://doi.org/10.29040/ijebar.v6i1.4394.

Antoni, R. S. (2025). Determinan Fraud pada Perusahaan Sektor Energi: Peran Komposisi Aset, Nature of Industry dan Intellectual Capital. *Journal of Accounting and Finance Management, 6*(4), 1956–1964. https://doi.org/10.38035/jafm.v6i4.2429.

Aviantara, R. (2023). Scoring the Financial Distress and the Financial Statement Fraud of Garuda Indonesia with «DDCC» as the Financial Solutions. *Journal of Modelling in Management, 18*(1), 1–16. https://doi.org/10.1108/JM2-01-2020-0017.

Azis, N. K., Madah Marzuki, M., Nik Abdul Majid, W. Z., & Rosman, R. (2025). Safeguarding the Value of Human Life from Maqasid Al-Shariah and Fraudulent Financial Reporting: Evidence from Malaysia's Shariah-Compliant Companies. *Journal of Islamic Accounting and Business Research.* https://doi.org/10.1108/JIABR-12-2023-0447.

Bansal, M. (2024). Earnings Management: A Three-Decade Analysis and Future Prospects. *Journal of Accounting Literature, 46*(4), 630–670. https://doi.org/10.1108/JAL-10-2022-0107.

Beneish, M. D. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal, 55*(5), 24–36.

Beneish, M. D., & Vorst, P. (2022). The Cost of Fraud Prediction Errors. *The Accounting Review, 97*(6), 91–121. https://doi.org/10.2308/TAR-2020-0068.

Biduri, S., & Tjahjadi, B. (2026). Determinants of Financial Statement Fraud: The Perspective of Pentagon Fraud Theory (Evidence on Islamic Banking Companies in Indonesia). *Journal of Islamic Accounting and Business Research, 17*(2), 392–422. https://doi.org/10.1108/JIABR-08-2022-0213.

Cahyono, S., Ardianto, A., & Sulaiman, N. A. (2025). From Dirty Money to "Creative Financing": An Exploratory Netnography, Euphemism and Cognitive Dissonance in Financial Crime Narratives. *Journal of Money Laundering Control, 28*(6). 701–726. https://doi.org/10.1108/JMLC-04-2025-0046.

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting Material Accounting Misstatements. *Contemporary Accounting Research, 28*(1), 17–82. https://doi.org/10.1111/j.1911-3846.2010.01041.x.

Demetriades, P., & Owusu-Agyei, S. (2022). Fraudulent Financial Reporting: An Application of Fraud Diamond to Toshiba's Accounting Scandal. *Journal of Financial Crime, 29*(2), 729–763. https://doi.org/10.1108/JFC-05-2021-0108.

Doss, S., & Bacha, S. (2026). Detecting Fraudulent Financial Statements through SCCORE Model: Evidence from French Context. *Journal of Financial Crime, 33*(1-2), 75–93. https://doi.org/10.1108/JFC-02-2025-0042.

du Toit, E. (2024). The Red Flags of Financial Statement Fraud: A Case Study. *Journal of Financial Crime, 31*(2), 311–321. https://doi.org/10.1108/JFC-02-2023-0028.

Ghorbel, N., & Ben Amar, A. (2025). Corporate Social Responsibility and Accounting Fraud: International Evidence. *International Journal of Ethics and Systems.* https://doi.org/10.1108/IJOES-09-2024-0290.

Haldar, A., Mayee, G. A., & Parikh, P. A. (2025). Stories Speak Louder than Numbers: Evidence from Annual Reports. *Qualitative Research in Financial Markets.* https://doi.org/10.1108/QRFM-10-2024-0286.

Halteh, K., & Tiwari, M. (2023). Preempting Fraud: A Financial Distress Prediction Perspective on Combating Financial Crime. *Journal of Money Laundering Control, 26*(6), 1194–1202. https://doi.org/10.1108/JMLC-01-2023-0013.

Hasan, M. S., Omar, N., Barnes, P., & Handley-Schachler, M. (2017). A Cross-Country Study on Manipulations in Financial Statements of Listed Companies: Evidence from Asia. *Journal of Financial Crime, 24*(4), 656–677. https://doi.org/10.1108/JFC-07-2016-0047.

Himmah, E. F., Nugaha, A., Putri, A., & Rahmani, A. N. (2025). Detecting of Fraudelent Financial Report: Perspective F Score Modeling. *InFestasi, 21*(2), 190-209. https://doi.org/10.21107/infestasi.v21i2.32246.

Husnurrosyidah, H., & Fatihah, I. (2022). Fraud Detecting Using M-Score and F-Score: Which is More Effective?. *Equilibrium: Jurnal Ekonomi Syariah, 10*(1), 137–151. http://dx.doi.org/10.21043/equilibrium.v10i1.15351.

Islamiyati, D., & Purnomo, D. E. (2021). Komparasi Akurasi Model M-Score dan Z-Score dalam Mendeteksi Fraudulent Financial Reporting. *Neraca, 17*(1), 1-18. https://doi.org/10.48144/neraca.v17i1.591.

Javaid, M. I., Sarwar, N., & Qureshi, M. U. (2025). Forecasting Corporate Earnings: Integrating Financial Ratios and Earnings Management Predictors. *Journal of Economic and Administrative Sciences.* https://doi.org/10.1108/JEAS-09-2023-0258.

Kaab Omeir, A., Vasiliauskaite, D., & Soleimanizadeh, E. (2023). Detection of Financial Statements Fraud Using Beneish and Dechow Models. *Journal of Governance and Regulation, 12*(3), 334-344. https://doi.org/10.22495/jgrv12i3siart15.

Attorney General's Office of the Republic of Indonesia. (2025). *Kerugian Negara dari Perkara Dugaan Korupsi Tata Kelola Minyak Mentah Pertamina Bertambah Jadi Rp285 Triliun.* Accessed at 4 October 2025 in https://story.kejaksaan.go.id.

Khamainy, A. H., Ali, M., & Setiawan, M. A. (2022). Detecting Financial Statement Fraud through New Fraud Diamond Model: The Case of Indonesia. *Journal of Financial Crime, 29*(3), 925–941. https://doi.org/10.1108/JFC-06-2021-0118.

Khatun, A., Ghosh, R., & Kabir, S. (2022). Earnings Manipulation Behavior in the Banking Industry of Bangladesh: The Strategical Implication of M-Score Model. *Arab Gulf Journal of Scientific Research, 40*(3), 302–328. https://doi.org/10.1108/AGJSR-03-2022-0001.

Kukreja, G., Gupta, S. M., Sarea, A. M., & Kumaraswamy, S. (2020). M-Score and Altman Z-Score as a Catalyst for Corporate Fraud Detection. *Journal of Investment Compliance, 21*(4), 231–241. https://doi.org/10.1108/JOIC-09-2020-0022.

MacCarthy, J. (2017). Using Altman Z-score and M-Score Models to Detect Financial Fraud and Corporate Failure: A Case Study of Enron Corporation. *International Journal of Finance and Accounting, 6*(6), 159-166. 10.5923/j.ijfa.20170606.01.

Madah Marzuki, M., Nik Abdul Majid, W. Z., Abu Bakar, H., Abdul Wahab, E. A., & Mohd Sanusi, Z. (2024). Risk Management Practices and Potential Fraudulent Financial Reporting: Evidence from Malaysia. *Asian Journal of Accounting Research, 9*(2), 116–126. https://doi.org/10.1108/AJAR-01-2022-0017.

Maniatis, A. (2022). Detecting the Probability of Financial Fraud Due to Earnings Manipulation in Companies Listed in Athens Stock Exchange Market. *Journal of Financial Crime, 29*(2), 603–619. https://doi.org/10.1108/JFC-04-2021-0083.

Masitah, D. (2024). Analisis Pengaruh Fraud Pentagon Terhadap Financial Statement Fraud Menggunakan M-Score dan F-Score. *Indonesian Journal of Innovation Studies, 25*(4). https://doi.org/10.21070/ijins.v25i4.1142.

Mulya, I. P., & Murjana, I. M., & Irianto, I (2024). Analisis Perbandingan dan Tingkat Akurasi Metode Altman Z-Score, Zmijewski, Springate dan Grover dalam Memprediksi Kebangkrutan Perusahaan. *Kredibel: Jurnal Ilmiah Manajemen, 3*(1), 269-279. https://doi.org/10.53512/kredibel.v3i1.496.

Ramzan, S., & Lokanan, M. (2025). The Application of Machine Learning to Study Fraud in the Accounting Literature. *Journal of Accounting Literature, 47*(3), 570–596. https://doi.org/10.1108/JAL-11-2022-0112.

Ratmono, D., Darsono, D., & Cahyonowati, N. (2020). Financial Statement Fraud Detection with M-Score and F-Score Model: An Empirical Analysis of Fraud Pentagon Theory in Indonesia. *International Journal of Financial Research, 11*(6), 154-164. https://doi.org/10.5430/ijfr.v11n6p154.

Repousis, S. (2016). Using Beneish Model to Detect Corporate Financial Statement Fraud in Greece. *Journal of Financial Crime, 23*(4), 1063–1073. https://doi.org/10.1108/JFC-11-2014-0055.

Sary, D., & Rani, P. (2025). Pengaruh Fraud Hexagon terhadap Pendeteksian Kecurangan Laporan Keuangan (Studi Empiris pada Perusahaan Sektor Energi yang Terdaftar di Bursa Efek Indonesia Periode 2019-2023). *Jurnal Review Pendidikan dan Pengajaran, 8*(2), 3867-3874.

Shahzadi, K., Alim, W., & Khan, S. N. (2025). Do the Fraud Triangle Components Fuel Complex Financial Fraud? A Study of Nonfinancial Firms in Pakistan. *Journal of Financial Crime, 32*(1), 207–220. https://doi.org/10.1108/JFC-10-2023-0270.

Singleton, T. W., & Singleton, A. J. (2010). *Fraud Auditing and Forensic Accounting.* John Wiley & Sons: Kanada.

Sitoresmi, D., Fakhruddin, I., Fitriati, A., & Setyadi, E. J. (2024). Pengaruh Fraud Pentagon Theory dalam Mendeteksi Fraudulent Financial Reporting dengan F-Score Model. *Owner: Riset dan Jurnal Akuntansi, 8*(4), 4341-4352. https://doi.org/10.33395/owner.v8i4.2294.

Soepriyanto, G., Tjokroaminoto, S., & Zudana, A. E. (2021). Annual Report Readability and Accounting Irregularities: Evidence from Public Listed Companies in Indonesia. *Journal of Financial Reporting and Accounting, 19*(5), 793–818. https://doi.org/10.1108/JFRA-01-2020-0006.

Suganda, D. A., Haura, R., Widodo, S., & Datu, Y. A. (2025). *Buku Referensi Metode Penelitian Kualitatif, Kuantitatif dan Campuran Teori dan Implementasi.* PT Media Penerbit Indonesia: Indonesia.

Sylwestrzak, M., & Słomkowska, J. (2025). Evaluating the Efficacy of Modified Beneish and Dechow Models for Detecting Earnings Management in Polish Companies. *Central European Management Journal.* https://doi.org/10.1108/CEMJ-07-2024-0225.

Tarjo, T., & Herawati, N. (2021). *Financial Fraud: Pendeteksian Menggunakan Aplikasi M-Score dan Data Mining.* Penerbit Adab.

Ye, X., Cheng, Z., Geng, X., & Zhu, C. (2022). Research on the Applicability of M-Score and F-Score Models to US-Listed Chinese Stocks. *In 2nd International Conference on Economic Development and Business Culture (ICEDBC 2022),* 329–341. https://doi.org/10.2991/978-94-6463-036-7_50.

Zhou, R., Yin, J., Feng, Y., Xu, C., & Zhu, R. (2022). Detection of Fraud Statement Using Calculation Models M-Score and F-Score: Evidence from Chinese Companies Listed in the United States. *In 2nd International Conference on Economic Development and Business Culture (ICEDBC 2022),* 780–790. https://doi.org/10.2991/978-94-6463-036-7_115.

***Coressponding Authors:**
Author can be contacted on E-mail: vinafebri911@gmail.com