Rapid Source Attribution of Illegal Crude Oil Dumping Using GC–FID Fingerprinting and Multivariate Analysis in the Rokan Field

Muhammad Hatta Nasution [a,1,*], Maulana Hardi [a,2], Adi Widiyanto [a,3], Ester Tio Minar E.Silalahi [a,4], Oktaviani Kusuma Wardani [a,5]

[a] Laboratorium Pertamina Hulu Rokan Minas, PT Pertamina Hulu Rokan, Riau, Indonesia
[1] atta_nst@yahoo.co.id *; [2] maul_94@yahoo.com; [3] adi.widiyanto@pertamina.com; [4] estertiominar@gmail.com; [5] oktavianikusumawardani99@gmail.com
*corresponding author

ARTICLE INFO

Article history
Received November 22, 2025
Revised December 24, 2025
Accepted December 29, 2025

Keywords
Analysis multivariate
Cosine similarity
GC–FID
Illegal dumping
Oil Fingerprinting

ABSTRACT

Illegal crude oil dumping constitutes a significant environmental offense that threatens ecosystem integrity and necessitates reliable forensic methodologies for accurate source attribution. This study proposes a rapid, practical, and cost-effective forensic framework that integrates routine gas chromatography–flame ionization detection (GC–FID) fingerprinting with multivariate statistical analysis using MALCOM software. The developed approach is intended to provide a scientifically defensible method for source correlation in operational oil fields, where timely identification of contamination sources is critically important. Soil and water samples were collected from two suspected dumping sites, designated Spot A and Mud Pit B, and compared with crude oil samples obtained from three reference wells: Wells 3D, 4D, and 1E. Source attribution was evaluated through comparative chromatographic fingerprint analysis, diagnostic biomarker ratios (Pr/Ph, Pr/n-C17, and Ph/n-C18), hierarchical clustering via dendrogram analysis, and cosine similarity to quantify chemical relationships among samples. The results revealed that the soil sample from Spot A exhibited an exceptionally high cosine similarity value (>0.999) and clustered closely with crude oil from Wells 3D and 4D, strongly indicating a common origin and supporting evidence of illegal discharge from these sources. In contrast, samples from Mud Pit B displayed distinct chromatographic characteristics and lower similarity values (<0.997), suggesting a different source and/or significant weathering. These findings demonstrate that integrating widely accessible GC–FID data with multivariate statistical tools offers a rapid, robust, and economically feasible approach for environmental forensic investigations and source attribution of illegal crude oil dumping incidents.

1. Introduction

Illegal crude oil dumping is a serious environmental violation that can contaminate soil, water, and surrounding ecosystems, particularly in regions with intensive oil and gas (O&G) operations. In active production areas such as the Rokan field, the extensive pipeline network, high fluid throughput, and frequent well intervention activities increase the likelihood of both accidental releases and intentional disposal events [1]. Accurate source identification is therefore critical for environmental monitoring and law enforcement, especially as recent advances in oil forensic methodology, such as the incorporation of polycyclic aromatic nitrogen heterocycle (PANH) biomarkers, have demonstrated improved diagnostic capability and higher resolution in differentiating crude oil sources [2].

Moreover, detailed characterization of the spilled oil is vital for determining appropriate remediation methods, as bioremediation strategies rely heavily on the chemical composition,

degradability, and microbial compatibility of hydrocarbons present in contaminated environments [3]. Complementing these approaches, rapid oil spill identification techniques using hydrophobic sampling paper combined with gas chromatography–mass spectrometry (GC/MS) have demonstrated high efficiency and analytical robustness, enabling more accurate and timely assessments in field-based investigations [4]. Oil fingerprinting has long been recognized as an effective approach for correlating oils based on their chemical characteristics. Conventional applications typically rely on compound distributions, such as n-alkanes, isoprenoids, and other hydrocarbon fractions, which are measurable using gas chromatography (GC) [5], [6], [7]. Beyond chromatographic and diagnostic ratio-based approaches, isotopic analyses have also proven effective in distinguishing hydrocarbon sources with high precision, particularly for polycyclic aromatic compound-bearing materials in environmental samples [8]. Advanced analytical frameworks that utilize biomarkers and comprehensive two-dimensional gas chromatography provide high-resolution differentiation of crude oils, offering deeper insights into compositional signatures relevant for forensic investigations [9].

Additionally, crude oil emulsion behavior and treatment efficiency depend heavily on the chemical nature of hydrocarbons present, as demonstrated by the development of effective local demulsifiers for oil-water separation in production systems [10]. Field surveillance technologies, such as free-floating PIPERS, further support environmental forensics by enabling the accurate localization of illegal tapping events along pipeline networks [11]. High-resolution biomarker analyses using GC-MS can further enhance specificity, but they require advanced instrumentation, longer processing times, and substantial analytical costs, limiting their use in rapid-response situations [12], [13]. Chromatographic gas analysis remains a foundational technique for hydrocarbon detection across diverse environmental and operational contexts [14]. Further analytical refinement has been achieved through comprehensive two-dimensional gas chromatography integrated with pixel-based chemometrics, enabling accurate multiparameter modeling of crude oil properties for advanced forensic interpretation [15]. In addition, detailed characterization of light hydrocarbons provides valuable geochemical insights into crude oil origins and reservoir behavior, strengthening source attribution and reservoir assessment efforts [16].

In contrast, GC–FID is widely available in routine production laboratories and can generate reproducible chromatographic profiles that reflect source-specific chemical patterns. Several studies have shown that GC–FID data, when supported by chemometric techniques such as hierarchical clustering or multivariate analysis, can effectively differentiate oils, including under weathered or altered conditions [13], [17], [18]. Analytical developments such as automated machine learning pipelines have improved the efficiency of geochemical data interpretation [19], while classical gas chromatography remains fundamental for detailed hydrocarbon fingerprinting [20]. Fingerprinting applications also extend to geological materials, where XRD clustering supports mineralogical discrimination [21], and spectroscopic methods offer high-resolution molecular insights into organic components in subsurface samples [22]. Ensuring analytical accuracy further requires adherence to standardized sampling guidelines such as those established by BSN [23]. However, the practical use of GC-FID as a rapid forensic tool remains underrepresented in the literature, especially regarding its integration with accessible analytical software for real-time decision support in operational oil fields.

While oil fingerprinting is well established, rapid-response applications in active oil fields often depend on complex, time-consuming analyses (e.g., GC-MS biomarker analyses). This study demonstrates that robust source differentiation can be achieved using standard GC-FID data routinely available in production laboratories, coupled with user-friendly multivariate software (Malcom). This approach provides a cost-effective, rapid forensic tool for field operators and regulators, as reported in the literature. Accordingly, this study aims to evaluate the application of GC-FID-based oil fingerprinting combined with multivariate analysis of the Malcom method for the rapid forensic identification of crude oil sources in a suspected illegal dumping incident in the Rokan field.

By presenting this case as a methodological application study, the research aims to demonstrate a practical workflow that enhances responsiveness and reliability in environmental forensic investigations within upstream O&G operations.

2. Research Methodology

This study employed an environmental forensic approach integrating GC-FID-based oil fingerprinting with multivariate statistical analysis. The workflow comprised field sampling, sample

preparation, chromatographic analysis, quality assurance/quality control (QA/QC), and chemometric evaluation.

2.1. Sampling and Site Context

Two suspected illegal dumping locations were investigated:

1) *Spot A (Soil)*
   a) *Located near a pipeline corridor.*
   b) *The material observed was a thin layer of oil-bearing soil.*
   c) *3 sub-samples collected from a 3 × 3 m grid area.*
   d) *Depth: 0–10 cm, composited into one representative sample.*
2) *Mud Pit B (Soil and Water)*
   a) *Historical reserve pit from past maintenance.*
   b) *The site contains standing water and moist sediments with hydrocarbon residue.*
   c) *Soil: 3 sub-samples collected around the pit margin, Depth: 0-10 cm, composited.*
   d) *Water: 2 sub-samples collected at opposite ends of the pit, Depth: surface to 10 cm, composited.*
3) *Reference Oils:*

Fresh crude oil samples from Wells 3D, 4D, and 1E (post-perforation) for comparison. Storage: amber vials at 4 °C.

2.2. Sample Preparation

Approximately 0.10 ± 0.01 g of each sample was transferred into a 40 mL vial and mixed with 2g anhydrous sodium sulfate (Na_2SO_4) to remove moisture. Samples were extracted using 10 mL dichloromethane (DCM) on a vortex mixer (1,800 rpm, 10 min). Extracts were filtered and transferred to GC vials for analysis.

2.3. GC-FID Instrumental Conditions

Analysis was performed using a Gas Chromatograph equipped with a Flame Ionization Detector (FID) using the following conditions for Table 1 below:

Table 1. Setting Performance GC-FID

Parameter	Setting
Instrument	Agilent/Shimadzu GC-FID (production lab standard)
Column	DB-5MS (30 m × 0.25 mm × 0.25 μm)
Carrier Gas	Helium, 1.0 mL/min constant flow
Injection Volume	0.1 μL
Injection Mode	Split 1:10
Inlet Temperature	300°C
Detector Temperature	320°C
Oven Program	60°C (2 min) → 6°C/min → 300°C (hold 25 min)
Total Run Time	110 min

These conditions provide clear resolution of n-alkanes and isoprenoids (Pr, Ph), enabling construction of diagnostic ratios and chromatographic fingerprints.

Note: biomarkers such as hopanes/steranes cannot be detected under GC-FID and are intentionally excluded

2.4. Quality Assurance and Quality Control (QA/QC)

The following procedures were applied:

a) *Procedural blanks: No detectable contamination observed.*
b) *Calibration standards (n-C13–n-C30 + isoprenoids): Retention time RSD < 0.5%.*
c) *Certified Reference Material (CRM–crude oil): Peak height reproducibility RSD < 3%.*
d) *Duplicate injections: RSD < 2%, confirming adequate precision.*
e) *Daily FID calibration: Conducted with standard hydrocarbon mixtures.*

QA/QC results verify that GC-FID performance was stable and meets environmental forensic requirements.

2.5. Chemometrics/Multivariate Analysis

Raw chromatograms imported to Malcom v2019.1.1.

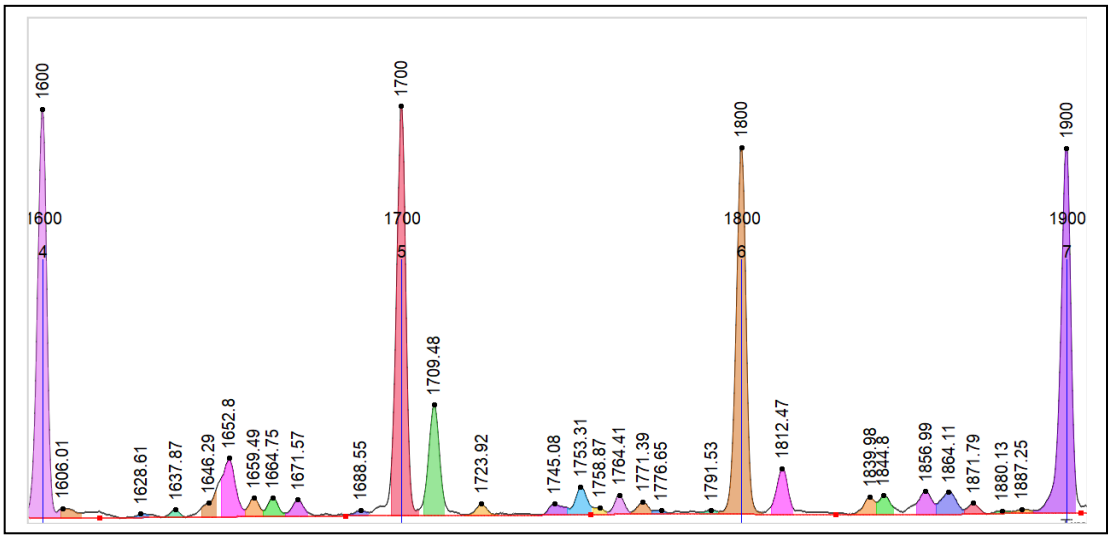

Fig. 1. Isomer peax indexing

1) Peak Indexing

a) n-alkanes were used as internal alignment markers.

b) Isomer window: 10 peaks left and right of the reference n-alkane.

2) Ratio Selection

a) Malcom computes numerous peak-height ratios.

b) The 25 ratios with the highest variance between samples (optimized for discrimination) were automatically selected.

3) Multivariate Analysis Methods

a) Distance Metric: Cosine similarity (0–1 scale).

b) Clustering Algorithm: Hierarchical Cluster Analysis (HCA).

c) Linkage Method: Ward's minimum variance.

d) Outputs Generated:

Dendrogram (cluster grouping)
Star Plot (pattern distribution)
Cosine similarity matrix
Diagnostic ratio plots (Pr/Ph, Pr/n-C17, Ph/n-C18)

These parameters ensure reproducibility and allow independent verification using R, Python, or SPSS.

3. Results and Discussion

3.1. Chromatographic Profiles and Initial Observations

GC–FID chromatograms (Fig. 2) reveal clear differences among the samples. Wells 3D and 4D display strong, well-resolved n-alkane distributions (n-C17, n-C18), whereas Well 1E exhibits lower overall signal intensity. The Spot A (soil) chromatogram shows pronounced n-alkane peaks and a general pattern that closely resembles Wells 3D/4D. In contrast, Mud Pit B (soil and water) exhibits reduced n-alkane intensity and subtle changes in the isoprenoid region (Pristane, Phytane), consistent with alteration/weathering relative to the reference oils.

The GC–FID profile of Spot A, similar to Wells 3D/4D, supports a common-source hypothesis. In contrast, the distinct pattern in Mud Pit B suggests a different origin or weathered oil rather than fresh crude from the perforated reference wells.

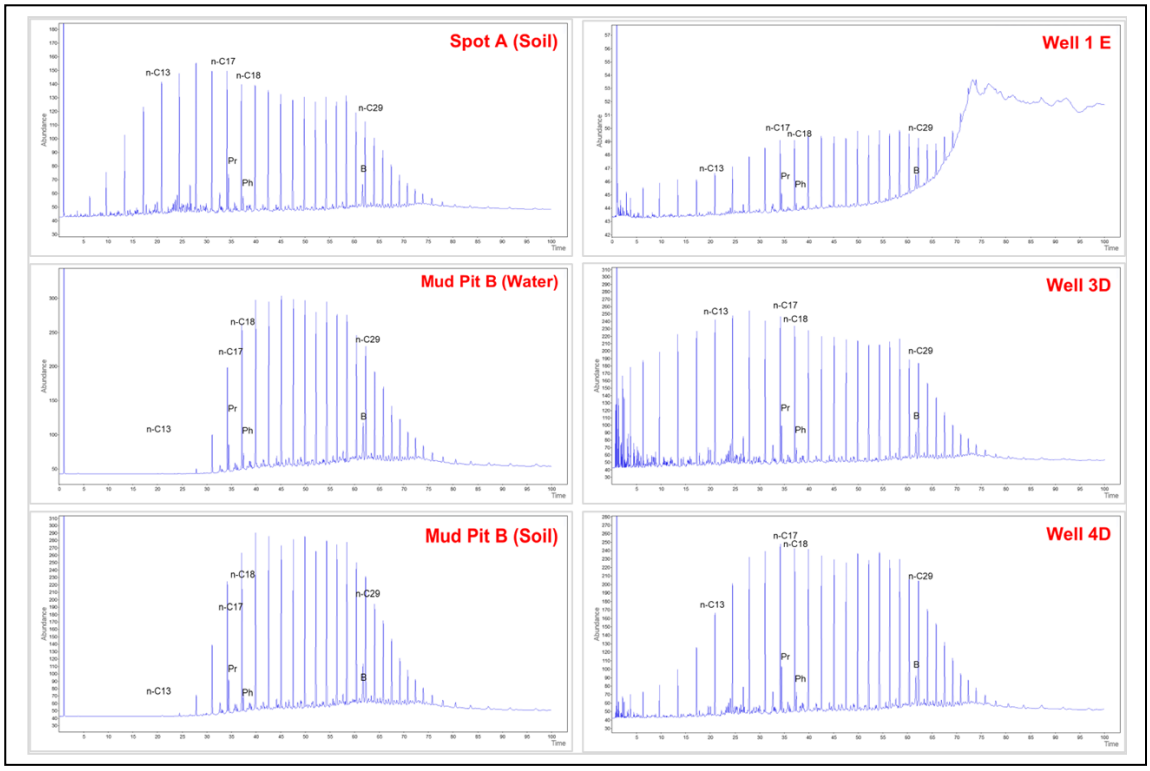

Fig. 2. Gas chromatogram results (GC-FID)

The presence of n-paraffins and isoprenoids indicates the potential existence of organic matter or hydrocarbons in the samples. Compounds such as Pristane and Phytane are commonly used in geochemical studies to identify organic sources and assess the degree of biodegradation. Botryococcane is a distinctive compound derived from the algae Botryococcus, which may suggest biological contributions from specific microorganisms [18]. The chromatogram results shown in Fig. 2 indicate that Well 1E contains a relatively low amount of crude oil, whereas Wells 3D and 4D exhibit higher crude oil content. Additionally, Spot A, represented by a soil sample, shows a notably high concentration of n-paraffin (n-C17). In contrast, the soil and water samples from Mud Pit B display a dominant presence of n-paraffin (n-C18).

3.2. Diagnostic Ratios and Evidence of Weathering

To strengthen interpretation, 3 diagnostic ratios, Pr/Ph, Pr/n-C17, and Ph/n-C18 were manually calculated from GC–FID peak heights (Table 4) and visualized as a Pr/Ph bar chart in Fig. 3. All samples show Pr/Ph > 1, indicating oxidative depositional conditions in line with environmental forensic precedent. Wells 3D/4D/1E occupy low to moderate ratio ranges (Pr/n-C17 = 0.26 - 0.28; Ph/n-C18 = 0.10 - 0.13), consistent with dominant n-alkanes and limited alteration. Spot A (soil) plots adjacent to the reference wells on the weathering plot (Pr/n-C17 = 0.26; Ph/n-C18 = 0.12), supporting source correlation. Mud Pit B (soil/water) shows Pr/n-C17 = 0.27; Ph/n-C18 = 0.12 with contrasting Pr/Ph (soil: 2.47; water: 1.52). Increased isoprenoid-to-n-alkane ratios suggest biodegradation (n-alkanes are more labile than isoprenoids). At the same time, the lower Pr/Ph in water could reflect water-washing/dilution, as hypothesized, rather than a definitive mechanism at this stage.

On the weathering plot (Pr/n-C17 vs. Ph/n-C18), fresh oils cluster at lower ratios; biodegradation tends to increase both ratios as n-alkanes decline relative to isoprenoids. The resulting spatial arrangement of samples corroborates the chromatographic observations above.

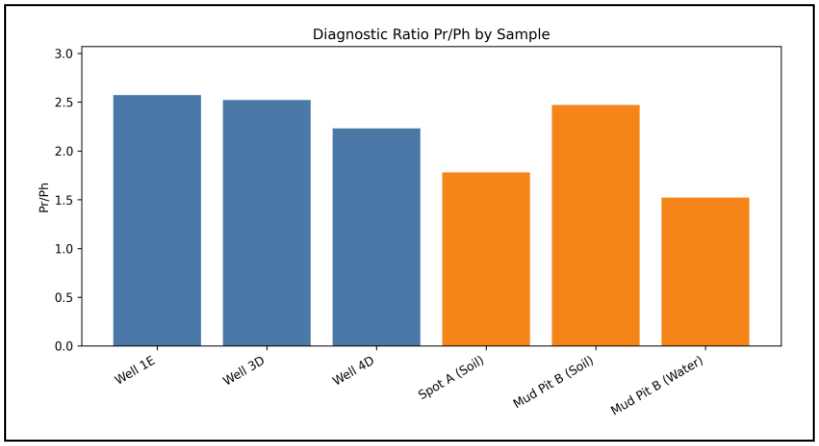

Fig. 3. Variation peak height in the Pr/Ph ratio graph

Fig. 3 illustrates the variation in Pr/Ph ratios across different locations, which can support geochemical interpretation in reservoir or environmental studies. All samples exhibit Pr/Ph ratios greater than 1, indicating that the environmental conditions during compound formation were predominantly oxidative. The highest values were observed in Well 1E (2.57) and Mud Pit B (Soil) (2.47), suggesting possible influences from biodegradation or differing organic inputs. Elevated ratio values reflect the oxidative potential of the environment and serve as important indicators in hydrocarbon quality assessment. This supports the conclusion that the samples are not derived from fresh crude oil and may have undergone biodegradation or originated from terrestrial organic sources.

Table 2. Manually Calculated Diagnostic Ratios

Sample	Pr/Ph	Pr/n-C17	Ph/n-C18
Well 1E	2.57	0.26	0.10
Well 3D	2.52	0.26	0.11
Well 4D	2.23	0.28	0.13
Spot A (Soil)	1.78	0.26	0.12
Mud Pit B (Soil)	2.47	0.27	0.12
Mud Pit B (Water)	1.52	0.27	0.12

3.3. Multivariate Statistical Correlation

We combined hierarchical clustering (dendrogram) with a cosine-similarity matrix in Table 3; a heatmap in Fig. 8. Cosine similarity indicates very high similarity between Spot A and Wells 3D/4D (> 0.999), with shared clustering on the dendrogram, providing strong evidence of a common source at GC FID level resolution. Mud Pit B (soil/water) shows lower similarity (<0.997, down to 0.979 vs. Well 3D for the water sample) and separate clustering, consistent with their divergent diagnostic ratios and distinct GC-FID profiles. Malcom's cosine similarity is computed from a set of discriminant peak ratios (e.g., "best 25 ratios"), not solely Pr/Ph; thus, it captures multivariate pattern similarity across chromatograms rather than single-ratio matches.

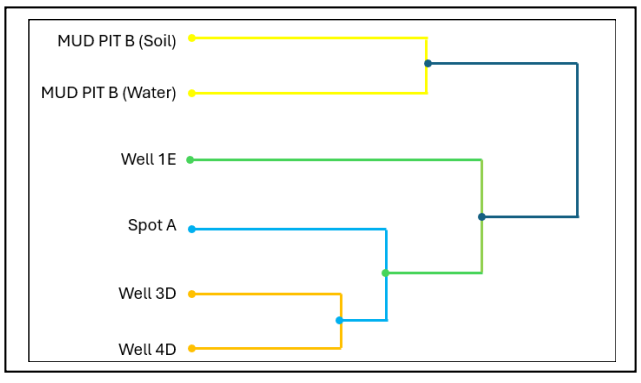

Fig. 4. Dendrogram diagram

Additionally, as shown in Fig. 4, the sample from Spot A exhibits a chromatographic pattern that closely resembles that of the crude oils from Well 4D and Well 3D, particularly in the distribution of n-alkanes and isoprenoids. This similarity is further supported by the dendrogram, which shows that all three samples are clustered together with a similarity level exceeding 90%. In contrast, the sample from Mud Pit B displays a significantly different chemical pattern compared to the three reference wells. The distribution of hopanes and steranes in this sample does not show a strong correlation with any of the wells, suggesting that the oil source in Mud Pit B likely originates from a different source not included among the reference wells.

3.4. Forensic Implications and Practical Application

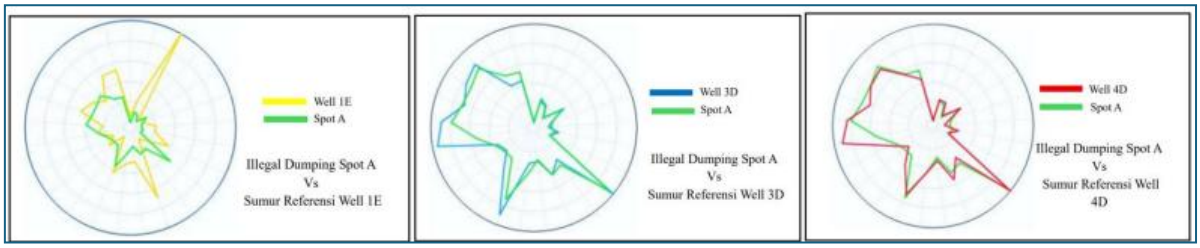

Fig. 5. Star diagram of illegal dumping spot a (soil) vs reference wells (well 1E, well 3D, and well 4D)

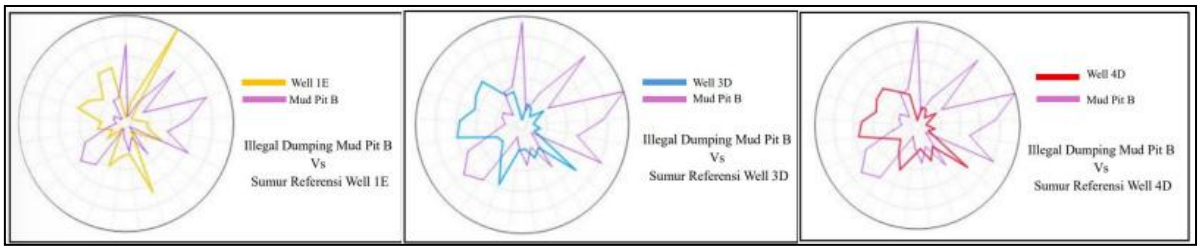

Fig. 6. Star diagram of illegal dumpling mud pit b (soil) vs reference wells (well 1E, well 3D, and well 4D)

Fig. 7. Star diagram of illegal dumpling mud pit b (water) vs reference wells (well 1E, well 3D, and well 4D).

Based on the star diagrams in Fig. 5, Fig. 6, and Fig. 7, it can be observed that the soil sample from Spot A and the reference wells (Well 1E, Well 3D, and Well 4D) are grouped within the same cluster and exhibit similar patterns to the crude oil from Well 4D and Well 3D. This strong similarity indicates that the sample from Spot A likely originates from one or both wells.

Table 3. Peak Height of Individual Compounds in Each Well (pA)

Compounds	n-C17	n-C18	Pristane	Phytane
Well 1E	5.33	5.32	1.39	0.54
Well 3D	199	184.42	51.41	20.38
Well 4D	200.78	194.85	56.83	25.44
Spot A (Soil)	179.66	216.1	47.31	26.58
Mud Pit B (Soil)	103.91	92.87	28.07	11.36
Mud Pit B (Water)	154.15	216.17	40.89	26.82

In contrast, the samples from Mud Pit B do not cluster with the reference wells and exhibit chemical patterns that differ from those of the crude oil from all three reference wells. This suggests

that the crude oil in the Mud Pit B samples is unlikely to be derived from fresh crude oil from wells that recently underwent POP operations, or that it has undergone degradation. These findings demonstrate that oil fingerprinting is an effective method for identifying the relationship between pollution sites and potential sources.

Table 4. Matriks Cosine Similarity

Sampel	Spot A (Soil)	Mud Pit B (Soil)	Mud Pit B (Water)	Well, 1E	Well, 4D	Well 3D
Spot A (Soil)	1	0.988925	0.97517	0.99837	0.99915	0.99976
Mud Pit B (Soil)	0.988925	1	0.99723	0.99548	0.99397	0.99151
Mud Pit B (Water)	0.975169	0.997233	1	0.98574	0.98316	0.97911
Well 1E	0.998366	0.995479	0.98574	1	0.99958	0.99932
Well 4D	0.999152	0.993969	0.98316	0.99958	1	0.99955
Well 3D	0.999758	0.991506	0.97911	0.99932	0.99955	1

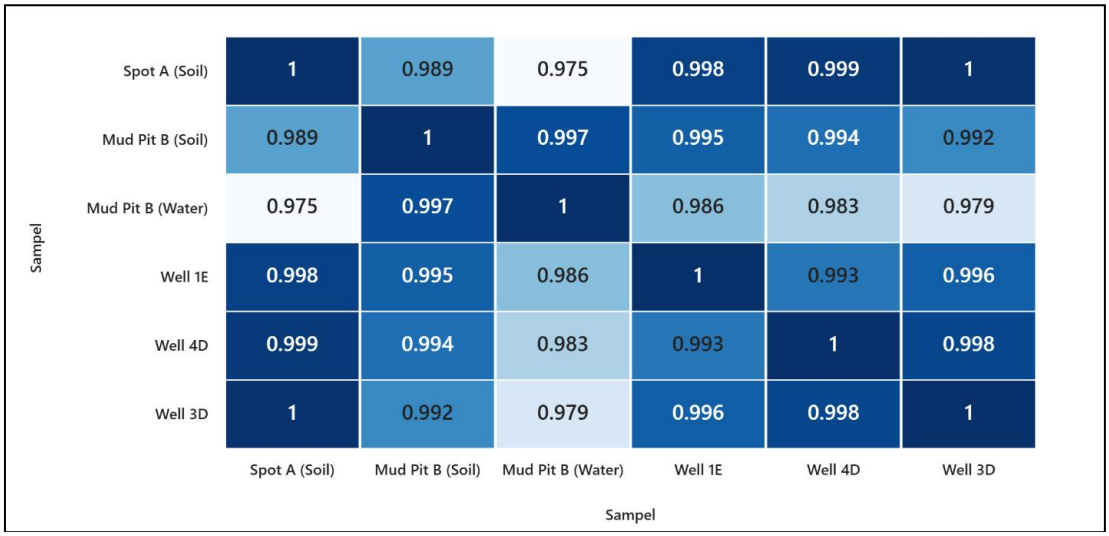

Fig. 8. Heatmap cosine similarity

The correlation values range from 0.9791 to 1.0000, indicating very strong relationships between locations. A perfect correlation of 1.0000 between Spot A, well 1E, and Well 4D suggests that the Pr/Ph ratios at these sites are nearly identical, reflecting similar compound distribution patterns. The correlation between Mud Pit B (Water) and other locations is slightly lower (e.g., 0.9791 with Well 3D), indicating minor differences in compound characteristics. A high correlation of 0.9970 between Mud Pit B (Soil) and Mud Pit B (Water) shows that, despite the different media (soil vs. water), the compound patterns remain highly similar. This heatmap demonstrates that, overall, Pr/Ph ratios across locations are highly correlated, suggesting similar compound sources, homogeneous geochemical processes, and potential stratigraphic or fluid migration connections between sites.

Weathering alters the chemical composition of spilled crude oil through evaporation, dissolution, photo-oxidation, and biodegradation [24]. These processes are evident in chromatograms and diagnostic ratios. In our study, GC-FID chromatograms of Spot A and reference wells (Well 3D and Well 4D) show strong n-alkane peaks with minimal flattening, indicating limited alteration. Conversely, Mud Pit B samples exhibit reduced n-alkane intensity and distortion in isoprenoid regions, consistent with biodegradation patterns [6], [9], [25]. All samples exhibit Pr/Ph > 1, confirming oxidative depositional conditions [7].

- Wells 1E and 3D show the highest Pr/Ph ratios (2.57 and 2.52), suggesting relatively fresh crude oil with minimal alteration.
- Spot A has a lower Pr/Ph ratio (1.78), but its Pr/n-C17 and Ph/n-C18 values remain close to those of the reference wells, supporting its linkage to Well 3D or Well 4D.
- Mud Pit B (Soil) shows a high Pr/Ph ratio (2.47), which may indicate selective loss of phytane or biodegradation.

- The water sample from Mud Pit B has the lowest Pr/Ph ratio (1.52), suggesting either advanced weathering or dilution.

Across all samples, Pr/n-C17 and Ph/n-C18 ratios remain below 0.3, confirming n-alkanes are still dominant, although slight weathering is evident in Mud Pit B, consistent with recent multivariate fingerprinting studies [26], [27]. These observations complement the dendrogram and cosine similarity outputs and reinforce the interpretation of source correlation and environmental alteration, as supported by advances in chemometric and machine-learning approaches for oil fingerprinting [28].

4. Conclusion

GC–FID chromatographic fingerprinting coupled with accessible multivariate analysis provides strong circumstantial evidence linking Spot A to Wells 3D/4D in this illegal dumping investigation. Mud Pit B samples are chemically distinct and may reflect a different source or a weathered counterpart. For definitive source attribution and genetic biomarker confirmation, targeted GC–MS analysis (e.g., hopanes, steranes) is recommended.

Acknowledgment

The authors would like to express their sincere gratitude to PT. Pertamina Hulu Rokan for providing access to laboratory facilities, sample materials, and technical support throughout the research process. The authors also acknowledge the contributions of the laboratory team in Minas, Riau, for their assistance in sample preparation and data analysis.

References

[1] N. Firdausiah, "Penegakan hukum pengeboran minyak ilegal pada pertambangan rakyat," *Constitution Journal*, vol. 1, no. 2, pp. 107–120, Dec 2022, doi: 10.35719/constitution.v1i2.18.

[2] P. McCallum, T. Filewood, J. Sawitsky, H. Kwok, P. Brunswick, J. Yan, L. Chibwe, K. Tikkisetty, and D. Shang, "Enhancement of oil forensic methodology through the addition of polycyclic aromatic nitrogen heterocycle biomarkers for diagnostic ratios," *Environ. Monit. Assess.*, vol. 195, no. 3, Mar 2023, doi: 10.1007/s10661-023-10941-3.

[3] M. I. Matilda and H. S. Samuel, "Bioremediation of oil spill: concept, methods and applications," *Discover Chemistry*, vol. 1, no. 1, Nov 2024, doi: 10.1007/s44371-024-00038-2.

[4] J. Shaw, J. Lesuk, L. Rabinovitch, T. Filewood, H. Kwok, J. Yan, P. Brunswick, T. Huan, and D. Shang, "Efficient oil spill identification utilizing hydrophobic sampling paper and gas chromatography/mass spectrometry," *Anal. Methods*, vol. 17, no. 29, pp. 6073–6087, Jul 2025, doi: 10.1039/D5AY00503E.

[5] M. T. Hidayat and E. Untoro, "Analisa penanggulangan dan pencegahan problem scale pada flowline sumur pmb-xxx pt pertamina hulu rokan prabumulih field," *Lembaran Publikasi Minyak dan Gas Bumi (LPMGB)*, vol. 44, no. 1, pp. 227–245, Dec 2022, doi: 10.29017/lpmgb.44.3.166.

[6] J. Sun, X. Wang, Q. Song, R. Li, J. Xie, X. Yang, L. Cai, Z. Wang, C. Zhao, and X. Zhang, "Fingerprint characteristics of refined oils and their traceability in the groundwater environment," *Chemosphere*, vol. 333, Aug 2023, doi: 10.1016/j.chemosphere.2023.138868.

[7] A. O. Aigberua, "Quantitative oil source fingerprinting and diagnostic ratios: Application for identification of soil residual hydrocarbon (SRH) in waste dump areas within oil well clusters," *Journal of Environmental Treatment Techniques*, vol. 7, no. 2, pp. 220–228, Jun 2019.

[8] J. M. E. Ahad, H. Pakdel, T. Labarre, C. A. Cooke, P. R. Gammon, and M. M. Savard, "Isotopic analyses fingerprint sources of polycyclic aromatic compound-bearing dust in athabasca oil sands region snowpack," *Environ. Sci. Technol.*, vol. 55, no. 9, pp. 5887–5897, May 2021, doi: 10.1021/acs.est.0c08339.

[9] B. M. Misselwitz, C. Jack, C. English, and B. Burger, "Fingerprinting crude oils and tarballs using biomarkers and comprehensive two-dimensional gas chromatography," *Restek Pure Chromatography*, pp. 1–14, 2014.

[10] C. Emuchuo and U. Osokogwu, "Novel local demulsifiers for crude oil emulsion treatment in oil and gas industry," *Petroleum Science and Engineering*, vol. 9, no. 2, pp. 48–54, Jul 2025, doi: 10.11648/j.pse.20250902.11.

[11] Z. Shand, A. Van Pol, and R. W. F. Dos Santos, "Detecting illegal tapping locations with free-floating pipers®," *Rio Oil and Gas Expo and Conference*, vol. 20, no. 2020, pp. 278–279, Dec 2020, doi: 10.48072/2525-7579.rog.2020.278.

[12] J. F. Huertas-pérez, C. C. Hernandez, A. N. Galindo, M. Dubois, L. Perring, A. Tarres, J. Nicolay, C. Vocat, and T. Delatour, "Discrepancies in mineral oil confirmation by two-dimensional gas chromatography – mass spectrometry: a call for harmonization," *Molecules*, vol. 30, no. 13, Jul 2025, doi: 10.3390/molecules30132830.

[13] S. Mitkidou, N. Kokkinos, E. Emmanouilidou, Y. Yohannah, T. Spanos, C. Chatzichristou, and A. Ene, "Investigation of petroleum hydrocarbon fingerprints of water and sediment samples of the Nestos River Estuary in Northern Greece," *Applied Sciences (Switzerland)*, vol. 12, no. 3, Feb 2022, doi: 10.3390/app12031636.

[14] R. Alamu, O. J. Dada, S. Limited, and B. John, "Chromatographic gas analysis for hydrocarbon detection," *ResearchGate*, May 2025.

[15] A. Moreira de Oliveira, C. A. Teixeira, J. E. B. Castiblanco, Í. Medeiros Júnior, and L. W. Hantao, "Multiparameter modeling for 10 crude oil properties using comprehensive two-dimensional gas chromatography and pixel-based chemometrics," *Energy and Fuels*, vol. 39, no. 27, pp. 12882–12892, Jul 2025, doi: 10.1021/acs.energyfuels.5c01554.

[16] D. M. Coutinho, C. L. Torres, V. B. Pereira, R. V. S. Silva, M. C. Santos, D. S. Dubois, J. P. Lopes, G. Vanini, F. R. A. Neto, and D. A. Azevedo, "Characterization of light hydrocarbons in brazilian pre-salt crude oils: applications in reservoir geochemistry," *Energy and Fuels*, vol. 39, no. 12, pp. 5789–5801, Mar 2025, doi: 10.1021/acs.energyfuels.4c06112.

[17] C. C. Chua, P. Brunswick, H. Kwok, J. Yan, D. Cuthbertson, G. V. Aggelen, and D. Shang, "Tiered approach to long-term weathered lubricating oil analysis: GC/FID, GC/MS diagnostic ratios, and multivariate statistics," *Anal. Methods*, vol. 12, no. 43, pp. 5236–5246, Oct 2020, doi: 10.1039/D0AY01510E.

[18] C. C. Chua, P. Brunswick, H. Kwok, J. Yan, D. Cuthbertson, G. V. Aggelen, C. C. Helbing, and D. Shang, "Enhanced analysis of weathered crude oils by gas chromatography-flame ionization detection, gas chromatography-mass spectrometry diagnostic ratios, and multivariate statistics," *J. Chromatogr.*, Dec 2020, doi: 10.1016/j.chroma.2020.461689.

[19] G. H. Alférez, O. A. Esteban, B. L. Clausen, and A. M. M. Ardila, "Automated machine learning pipeline for geochemical analysis," *Earth Sci. Inform.*, vol. 15, no. 3, pp. 1683–1698, Sep 2022, doi: 10.1007/s12145-022-00821-8.

[20] R.Desrina, "Metode kromatografi gas untuk fingerprinting tumpahan minyak bumi di perairan," *LPMBG*, vol. 44, no. 2, pp. 117–128, Aug 2010, doi: 10.29017/LPMGB.44.2.158.

[21] S. Makvandi, M. Panalytical, J. Cauzid, A. Tarantola, V. Melfos, and P. Voudouris, "XRD clustering and quantitative analysis as a fingerprinting technique to study ore deposits," *The 17th SGA Biennial Meeting*, Aug 2023.

[22] D. Mihrin, M. Li, M. Alejandra Sánchez, C. Hoeck, R. W. Larsen, and K. L. Feilberg, "Spectroscopic fingerprinting of organic material extracted from tight chalk core samples of the North Sea," *ACS Omega*, vol. 5, no. 49, pp. 31753–31764, Dec 2020, doi: 10.1021/acsomega.0c04431.

[23] B. S. Nasional, "Metoda pengambilan contoh air permukaan," *Republik Indonesia*, vol. 59, no. 2, pp. 19, 2008.

[24] Z. Wang and M. F. Fingas, "Development of oil hydrocarbon fingerprinting and identification techniques," *Mar. Pollut. Bull.*, vol. 47, no. 9–12, pp. 423–452, Sep 2003, doi: 10.1016/S0025-326X (03)00215-7.

[25] F. D. Hostettler, Y. Wang, Y. Huang, W. Cao, B. A. Bekins, C. Rostad, C. F. Kulpa, and A. Lauren, "Forensic fingerprinting of oil-spill hydrocarbons in a methanogenic environment," *Environ. Forensics*, vol. 8, no. 1–2, pp. 139–153, Mar 2007, doi: 10.1080/15275920601180685.

[26] C. Sandu, A. C. Silver, and P. Srinivasan, "Enhancing resolution of geochemical fingerprinting by significant parameters selection for reservoir compartmentalization evaluation using unsupervised machine learning," in *AAPG/OnePetro Conference Proceedings*, 2025.

[27] M. Alferez, O. A. Esteban, B. Clausen, and A. Martinez, "Automated machine learning pipeline for geochemical analysis," *Earth Sci. Inform.*, vol. 14, no. 3, pp. 1123–1138, Sep 2021, doi: 10.1007/s12145-021-00645-7.

[28] Y. Zhang, L. Zhang, Z. Lei, F. Xiao, Y. Zhou, and X. Qian, "Unsupervised machine learning-based singularity models for geochemical anomaly identification," *Fractal and Fractional*, vol. 8, no. 2, pp. 112, Sep 2024, doi: 10.3390/fractalfract8020112.