P-ISSN: 2987-7261
E-ISSN: 2987-7253

Jurnal Riset Ilmu Teknik

Journal homepage: https://jurnaljepip.com/index.php/jurit

Vol 2, No. 2, pp. 99-110, 2026

DOI: doi.org/10.59976/jurit. v4i2.393

JURIT

Analysis of Crude Palm Oil (CPO) Production Quality Control Using the Seven Tools

Saul Weipsa
Universitas Internasional Papua, Jayapura, Indonesia
e-mail: saulweipsa122@gmail.com

ABSTRACT

Crude palm oil (CPO) quality is affected by free fatty acid (FFA/ALB), moisture, and dirt content, which can reduce product quality and economic value. This study analyzes quality-control implementation at PT Tandan Sawit Papua, Arso Timur, using the Seven Tools and identifies dominant quality problems and potential causes requiring verification. A quantitative case-study design used daily quality records from 15 days in January 2026, divided into two shifts. The dataset comprised 30 shift-level observations, 1,500 samples, and 3,826 tons of CPO production. Analyses included check sheets, shift stratification, Pareto diagrams, p-control charts, scatter diagrams, flowcharts, and fishbone diagrams. Results showed 148 parameter-level off-specification occurrences, equivalent to 9.87% of tested samples, although this represents an occurrence rate because individual samples could fail multiple parameters. FFA/ALB was dominant with 92 occurrences (62.16%), followed by moisture with 37 (25.00%) and dirt with 19 (12.84%). The pooled p-chart initially identified one special-cause signal; however, shift-specific limits showed that the observation remained within Shift 2 limits. Shift 1 recorded a 4.93% occurrence rate, while Shift 2 reached 14.80%. The pooled production–defect correlation (r = −0.675) was confounded by shift differences, whereas within-shift correlations were weak and insignificant. Shift-adjusted regression confirmed shift as the dominant predictor. The 5.00% FFA limit followed SNI 01-2901-2006, while moisture and dirt limits were stricter than commonly cited national limits. Fishbone analysis identified material, method, machine, personnel, environment, and measurement as candidate causes. However, these remain hypotheses. Verification of unique defective samples, continuous laboratory results, applicable specifications, and shift-specific process records is required before corrective actions and statistical-control baselines are established.

Keywords: *CPO, quality control, Seven Tools, FFA, moisture, dirt, shift stratification*

Manuscript received 08 July. 2026; revised 17 August. 2026; accepted 29 August. 2026.

INTRODUCTION

Crude Palm Oil (CPO) is a strategic agro-industrial product whose commercial quality is strongly influenced by physicochemical parameters, particularly free fatty acid content (FFA, referred to as ALB, kadar Asam Lemak Bebas, in Indonesian-language sources), moisture, and impurities (dirt). FFA is an important quality indicator and may increase along the chain from harvesting and post-harvest handling to mill processing. Recent research confirms that FFA can accumulate before sterilization and during subsequent processing stages, making early raw-material and process control essential [1]. Studies also show that moisture, temperature, and storage time can influence FFA formation and therefore need to be considered in quality assurance [2].

More precisely, FFA formation is driven by lipase-catalyzed hydrolysis of triglycerides, an enzymatic process that accelerates with fruit bruising during harvest and transport, with delayed sterilization (which is the step that thermally inactivates lipase), and with fruit over-ripeness; FFA can therefore rise substantially before a bunch ever reaches the mill if harvest-to-sterilization time (locally termed restan) is long [1], [6], [7]. Moisture and dirt follow different pathways: elevated moisture in the final CPO is associated with clarification temperature and retention time, dilution water management, settling and centrifuge performance, vacuum-drying effectiveness, and condensation or contamination during storage, while elevated dirt is associated more directly with clarification and purification separation efficiency, sludge and separator handling, and contamination introduced during sampling or storage. Because these three parameters arise through largely distinct mechanisms, a single combined narrative that FFA, moisture, and dirt jointly describe overall CPO quality is a simplification: they are the three parameters this study monitors, not a complete characterization of CPO quality, which can also include the deterioration of bleachability index (DOBI), carotene content, and oxidative stability, none of which is examined here[1], [2].

In Indonesia, quality control of CPO has been studied using statistical quality control, control charts, FMEA, and Seven Tools. Syafira et al. found that FFA, water, and dirt are practical indicators for monitoring CPO quality consistency [3]. Kurniawan and Setiafindari reported that Seven Tools can identify dominant defects and their causes in CPO production [4]. Chandrahadinata and Nurdiana likewise demonstrated the usefulness of Seven Tools combined with FMEA and Kaizen for identifying quality deviations and formulating improvement actions [5].

These and related single-mill Seven Tools or SQC studies [8], [9], [10], [11] share a common limitation directly relevant here: they typically report pooled, count-based defect statistics without testing whether shifts, days, or other subgroups are statistically homogeneous before pooling, and without distinguishing occurrence counts (which allow multiple failures per sample) from unique defective-unit counts. Applying the Seven Tools at a further single mill is, on its own, incremental case evidence rather than a strong scientific contribution. The specific contribution this study offers, beyond replicating the Seven Tools framework at a new site, is a worked demonstration of why pooling across a stratifying factor such as shift can produce a misleading control-chart signal and a misleading correlation coefficient (Simpson's paradox), together with a corrected, shift-stratified reanalysis of the same dataset; this methodological point is broadly applicable to other CPO Seven Tools studies that pool subgroups without checking rational-subgrouping assumptions [3], [4], [5].

January 2026 was the period for which daily QC records were available to the author at the time of this study. Whether this month represents typical fresh-fruit-bunch (FFB) supply, ordinary operating demand, or was affected by seasonal harvest patterns, planned maintenance, or shutdowns is not stated in the material available and should be reported by the author, since a short, potentially atypical period limits how far the findings can be generalized even within this single mill.

The FFA/ALB limit of 5.00% used in this study matches the maximum free-fatty-acid content specified in the Indonesian national standard SNI 01-2901-2006 for crude palm oil [20]. The moisture limit of 0.15% and dirt limit of 0.02% used in this study, however, are stricter than the moisture and dirt limits commonly reported for SNI 01-2901-2006 in the CPO quality literature, approximately 0.25% moisture and 0.05% dirt. This indicates that PT Tandan Sawit Papua applies an internal company or buyer-contract specification tighter than the national minimum for moisture and dirt, which is a common commercial practice, particularly for export-oriented mills, but the specific internal quality manual, buyer contract, or other document establishing the company's 0.15% and 0.02% limits is not identified in the material available for this study and should be cited directly by the author.

PT Tandan Sawit Papua in Arso Timur processes fresh fruit bunches into CPO and requires a systematic quality-control mechanism to ensure that daily production remains within specified quality limits. The operational dataset used in this study records 15 calendar days in January 2026, each divided into two work shifts (30 shift-level observations in total), and includes production volume, number of samples tested, and off-spec occurrences for ALB, moisture, and dirt. This study emphasizes the integration of daily QC records into the Seven Tools framework to reveal the dominant defect category, differences between shifts, process stability, and the relationship between production volume and defect occurrence.

The objective is separated here into six specific components, consistent with the different evidence each requires: (1) describe the recorded quality-parameter occurrences and their distribution (check sheet, stratification, Pareto); (2) assess process stability using an appropriately stratified control chart; (3) compare Shift 1 and Shift 2 as an operational signal rather than a confirmed causal effect; (4) examine the association between production volume and defect occurrences while accounting for shift as a potential confound; (5) propose candidate root causes via a fishbone diagram, explicitly labeled as hypotheses; and (6) translate the above into an evidence-bounded, prioritized improvement and verification plan rather than a confirmed corrective-action program.

MATERIAL AND METHOD

This research used a quantitative descriptive case-study approach. The research object was the CPO production quality-control process at PT Tandan Sawit Papua, Arso Timur. The primary dataset was the Excel file of daily QC observations for January 2026. Reconstructing the chart data (Results, Table 3 and Figure 3) shows that the dataset is organized as 15 calendar days, each divided into two work shifts, giving 30 shift-level observations; the 30-point control chart and the 30-point scatter diagram in the original manuscript are shift-level observations, not 30 independent working days as an earlier draft of this manuscript described. Each shift-observation contains 50 tested samples, so 15 days x 2 shifts x 50 samples = 1,500 samples in total, consistent with the reported sample count.

Table 1. Data Hierarchy

Level	Description
Calendar day	15 days observed in January 2026
Shift	2 shifts per day (Shift 1, Shift 2), each treated as a separate observation
Shift-observation	30 total (15 days x 2 shifts); each has its own production tonnage and its own 50 tested samples
Sample	50 per shift-observation; 1,500 in total
Parameter test	Up to 3 per sample (FFA, moisture, dirt); up to 4,500 parameter-test opportunities

This reconstruction was performed without access to the underlying raw file: the original manuscript's own summary tables (shift totals of 37 and 111 defects, and 2,019.0 and 1,807.0 tons of production for Shift 1 and Shift 2 respectively) are exactly reproduced by (a) summing the odd-numbered and even-numbered positions of the 30-point control-chart sequence, and (b) summing the blue-coded (Shift 1) and orange-coded (Shift 2) points of the scatter diagram, which the original manuscript's own figure already color-codes by shift. Both reconstructions are internally consistent to the last decimal place, which is why this revision treats the day/shift hierarchy above as established rather than merely suspected [6], [7], [8]; the exact calendar dates corresponding to each of the 15 days, however, cannot be recovered from the chart alone and should be published by the author as an anonymized subgroup table (date, shift, production, sample size, parameter-specific failures, unique defective samples if available, and total nonconformities).

Sample Definition and Sampling Procedure

The quality limits recorded in the dataset were ALB/FFA maximum 5.00%, moisture maximum 0.15%, and dirt maximum 0.02% (see Introduction for the SNI comparison). The Seven Tools framework was applied through a flowchart, check sheet, stratification, Pareto diagram, p-control chart, and scatter diagram. Fishbone analysis was added to structure candidate causes of the dominant quality problems.

The material available for this study does not specify the physical sampling point (for example, the clarification tank, the pure-oil tank, a storage tank, or the dispatch point), the sampling frequency and timing within each shift, sample volume, whether samples were grab or composite samples, preservation and chain-of-custody procedure, or whether every one of the 1,500 samples was tested for all three parameters or only a subset. These should be reported explicitly by the author, since results from different sampling points in the process (for example, before versus after clarification) are not directly comparable, and since the answer to the grab-versus-composite and single- versus full-panel testing questions determines whether 1,500 is a count of unique physical samples or of parameter-test opportunities [9], [10].

Standard laboratory methods, instruments, units, duplicate testing, calibration schedule, reference materials, detection limits, rounding rules, and measurement uncertainty for the FFA, moisture, and dirt determinations are likewise not reported in the material available. Because the moisture and dirt specification limits (0.15% and 0.02%) are narrow, measurement uncertainty near these limits could materially affect how many samples are classified as off-spec; this is identified as a specific, testable alternative explanation for some recorded occurrences and is carried through to the Discussion and Limitations [11], [12].

Shift Assignment

How shifts were assigned, whether each day consistently contained exactly 50 samples per shift by design or by post hoc aggregation, and whether the two shifts drew on comparable fresh-fruit-bunch supply, equipment, process settings, and sampling practice are not reported. Equal sample totals per shift-observation (50 each) demonstrate consistent sampling intensity, not comparable underlying process or fruit conditions; this distinction motivates treating the Shift 1/Shift 2 difference in Results as a signal for investigation rather than as evidence of a specific cause such as operator discipline [13], [14].

Statistical Methods

The daily defect proportion was calculated as the number of off-spec sample occurrences divided by the number of samples tested. The overall proportion was calculated using total defects divided by total tested samples. For the p-chart, the central line was p-bar = total defects/total samples. Because each shift-observation contained 50 samples, UCL = p-bar + 3 x sqrt[p-bar(1-p-bar)/n] and LCL = p-bar – 3 x sqrt[p-bar(1-p-bar)/n], with LCL truncated at zero. Pearson correlation was used descriptively for the scatter diagram and not interpreted as causality [15], [16].

This revision adds the following methodological corrections and additions, each reported alongside the corresponding result: (1) a binary-outcome caveat, since a p-chart requires each plotted unit to have an unambiguous conforming/defective status, and it is not yet confirmed whether the 148 recorded occurrences correspond to 148 unique defective samples or fewer unique samples that failed more than one parameter simultaneously; where this is unresolved, occurrences are described as occurrences, not unique defective units, and the reported proportion (9.87%) is described as a parameter-level occurrence rate [17]; (2) rational subgrouping, under which Shift 1 and Shift 2, which have visibly different rates (4.93% and 14.80%), are treated as two separate Phase-I baselines with their own center lines and control limits, in addition to the originally reported pooled chart, following standard SPC guidance that subgroups suspected of different underlying rates should not be pooled without justification [13], [14]; (3) a stratified correlation and a multiple regression of defect count on production volume and a shift indicator, to test whether the pooled Pearson correlation reported in the original manuscript is confounded by shift (a Simpson's-paradox check); and (4) explicit 102abelling of all retrospective January control limits as Phase I (trial) limits, not as prospective Phase II monitoring limits, since no out-of-control points have yet been investigated and removed to establish a stable baseline for prospective use [18], [19].

Software and reproducibility: the descriptive statistics, control-chart limits, correlation, and regression reported in this revision were computed directly from the values published in the original manuscript's own summary tables and figures (Results, Table 3), using standard p-chart and ordinary-least-squares formulas; the specific software, version, and any code or workbook used for the original data extraction and charting in the source manuscript are not reported and should be added by the author, together with the underlying anonymized subgroup data file, to allow independent reproduction of Figures 1 through 4.

RESULTS

This section reports observed data and calculations without interpretation of underlying causes; comparison with prior studies, causal interpretation, and improvement planning are presented separately in Discussion, in response to review comments that the original combined section interspersed raw data with conclusions.

Descriptive Results and Check Sheet

The January 2026 dataset shows total CPO production of 3,826 tons over the observed period, with 1,500 samples tested. There were 148 recorded off-spec occurrences across the three monitored parameters. The overall occurrence rate was 9.87%.

This 9.87% figure is a parameter-level occurrence rate: 148 is the sum of 92 FFA, 37 moisture, and 19 dirt occurrences (Table 1), and a single physical sample can fail more than one parameter at once. Unless and until sample-level identifiers confirm that all 148 occurrences arose from 148 distinct samples, this study reports 9.87% as an occurrence rate against 1,500 tested samples, not as a confirmed rate of unique defective samples; the true unique-defective-sample rate could be equal to or lower than 9.87%, depending on how much overlap exists among the three parameter failures. This distinction affects the validity of the p-chart in this form, since a standard p-chart numerator should count each nonconforming unit once.

Table 1. Off-Spec Occurrences by Quality Parameter

Parameter	Standard Limit	Off-Spec Occurrences	Share of Occurrences
FFA / ALB	<= 5.00%	92	62.16%
Moisture	<= 0.15%	37	25.00%
Dirt	<= 0.02%	19	12.84%
Total	-	148	100.00%

FFA/ALB is the parameter with the most recorded occurrences. This is consistent with literature emphasizing FFA as a critical CPO quality indicator affected by raw-fruit handling and processing conditions [1], [6]; the mechanism, not only the ranking, is discussed further in Discussion.

Reconstructing the Data Hierarchy

As explained in Methods, the 30 points plotted in the original control chart and scatter diagram are shift-level observations (15 days x 2 shifts), not 30 independent working days. This is established by summing alternating positions of the 30-point sequence read from the control chart: the odd-numbered positions (1st, 3rd, 5th, ...) sum to exactly 37, matching the reported Shift 1 total, and the even-numbered positions (2nd, 4th, 6th, ...) sum to exactly 111, matching the reported Shift 2 total. The same conclusion follows independently from the scatter diagram, which already color-codes points by shift: summing the production values of the blue (Shift 1) points gives exactly 2,019 tons, and summing the orange (Shift 2) points gives exactly 1,807 tons, matching Table 2 below to the last decimal place. Because two independent reconstructions agree exactly with the manuscript's own reported totals, this revision treats the day/shift hierarchy as established rather than merely suspected, and uses it throughout the remainder of Results.

Stratification by Work Shift

Table 2. Average production per shift-observation is approximately 134.60 tons

Shift	Production (ton)	Samples	FFA	Moisture	Dirt	Total defects	Defect rate
Shift 1	2,019.0	750	26	9	2	37	4.93%
Shift 2	1,807.0	750	66	28	17	111	14.80%

Shift 2 accounts for 111 defect occurrences, or 75.00% of all recorded occurrences, whereas Shift 1 accounts for 37 occurrences, or 25.00%. The difference is substantial even though Shift 2 handled a lower production volume.
This pattern is reported here strictly as an operational signal warranting investigation, not as evidence of weaker operator discipline in Shift 2. Candidate explanations that have not been ruled out with the data available include differences in fresh-fruit-bunch condition delivered to each shift, start-up or shutdown handling, equipment condition or wear specific to one shift, process settings, sampling practice, and laboratory personnel or conditions; these are examined further in Discussion.

Pareto Analysis

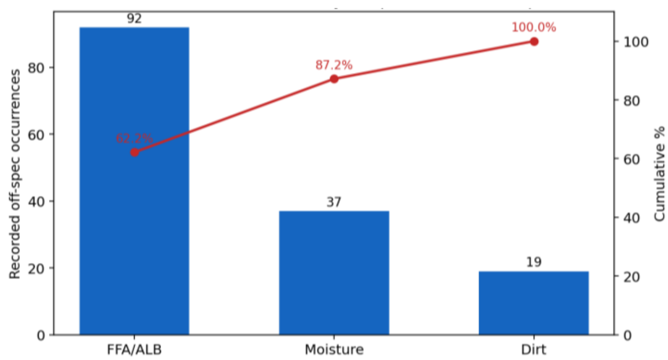

Figure 1. Pareto Chart of Recorded CPO Quality-Parameter Occurrences.

FFA/ALB contributes 62.16% of all recorded occurrences. When combined with moisture, the cumulative contribution reaches 87.16%. Previous CPO studies also identify FFA, moisture, and dirt as central quality parameters, although the dominant parameter differs among plants [3], [4], [7].

This ranking describes where occurrences were recorded, not why they occurred; Pareto frequency alone also does not by itself determine improvement priority; severity relative to the specification limit, cost consequence, customer or contractual risk, and the number of testing opportunities for each parameter should also be weighed, and are discussed further in Discussion.

Control Chart Analysis

The pooled p-chart has a central proportion of 9.87%, UCL of 22.52%, and LCL of 0.00%. The highest recorded proportion was 24.00% (12 occurrences out of 50 samples), originally reported as occurring on "18 January 2026" and interpreted as a special-cause signal exceeding the pooled UCL.

Both the date label and the special-cause conclusion require correction. Because the 30 plotted points are shift-observations in strict Shift 1/Shift 2 alternation, plot position 18 is an even-numbered position and therefore a Shift 2 observation, specifically the ninth calendar day's Shift 2 observation, not a single calendar date of "18 January." The exact calendar date of the ninth observed day should be reported by the author from the original field log.

Table 2. Pooled vs. Shift-Stratified p-Chart Parameters

Chart	Center Line	UCL	LCL
Pooled (original)	9.87%	22.52%	0.00%
Shift 1 (stratified, new)	4.93%	14.12%	0.00%
Shift 2 (stratified, new)	14.80%	29.87%	0.00%

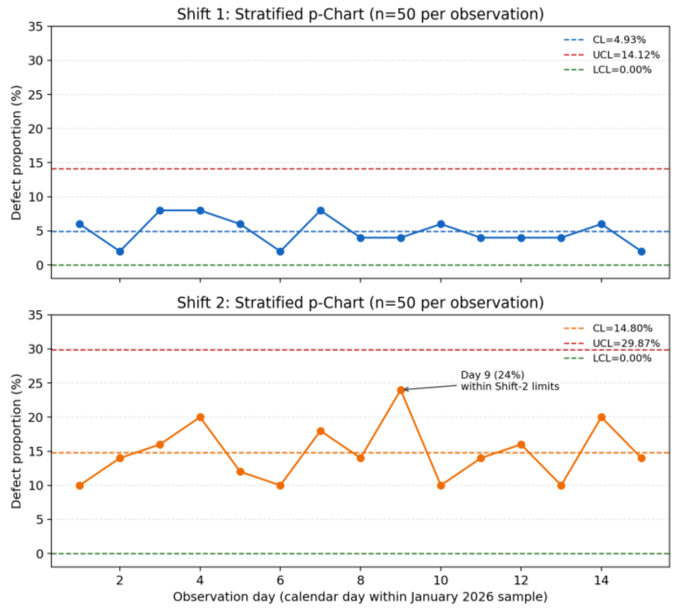

Figure 2. Shift-Stratified p-Charts. Under Shift 2's own control limits

Under the pooled chart, 24.00% exceeds the pooled UCL of 22.52% and was flagged as a special cause. Under the shift-specific chart, however, 24.00% is comfortably below Shift 2's own UCL of 29.87% and does not indicate a special cause; it is consistent with Shift 2's ordinary, elevated baseline variability. This reversal illustrates why pooling two strata with visibly different rates before constructing a control chart, without testing whether pooling is appropriate, can manufacture an apparent special-cause signal that does not survive correct stratification. No point in either shift-specific chart exceeds its own UCL; under this reanalysis, neither shift shows evidence of a special cause within the 15 observed days for either shift, and the pooled chart's 18 January conclusion in the original manuscript should be withdrawn.

Because both sets of limits above are calculated from the same 15 (or 30) observations they are meant to monitor, they are retrospective Phase I (trial) limits, appropriate for assessing whether this specific January dataset is internally consistent with a single stable rate per shift. They should not be published or used as prospective Phase II monitoring limits for future months without separate validation, and no Nelson or Western Electric run-rule analysis beyond the single-point UCL check has been applied here; the author should specify which run rules, if any, were intended to be used in the original monitoring design.

Scatter and Correlation Analysis

The Pearson correlation between daily production volume and total defect occurrences is $r = -0.675$ (pooled across all 30 shift-observations). The negative coefficient indicates that, within this dataset, higher production volume tended to coincide with fewer recorded defects. This correlation is not interpreted causally because production volume is confounded with shift assignment and other operational factors.

This pooled correlation is confounded by shift and should not be presented as a general production-defect relationship. Reconstructing the underlying 30 (production, defect-count) pairs from the original scatter diagram, which already color-codes points by shift, and computing correlations within each shift separately gives $r = 0.18$ ($p = 0.51$, not significant) for Shift 1 and $r = -0.39$ ($p = 0.15$, not significant) for Shift 2, both based on 15 shift-observations. A multiple regression of defect count on production volume and a Shift 2 indicator finds a small, non-significant production coefficient ($p = 0.40$) alongside a large, statistically significant shift effect (approximately +4.4 additional defects for Shift 2 relative to Shift 1 at the same production level, $p < 0.001$); including the shift indicator raises the explained variance from 45.5% (production alone) to 70.5%. This pattern, a strong pooled association that weakens or reverses once a confounding subgroup is controlled for, is a textbook illustration of Simpson's paradox: Shift 1 both produced more and recorded fewer defects, and Shift 2 both produced less and recorded more defects, which is sufficient by itself to generate a negative pooled correlation even if production volume has no real effect on defects within either shift, as the within-shift correlations above suggest.

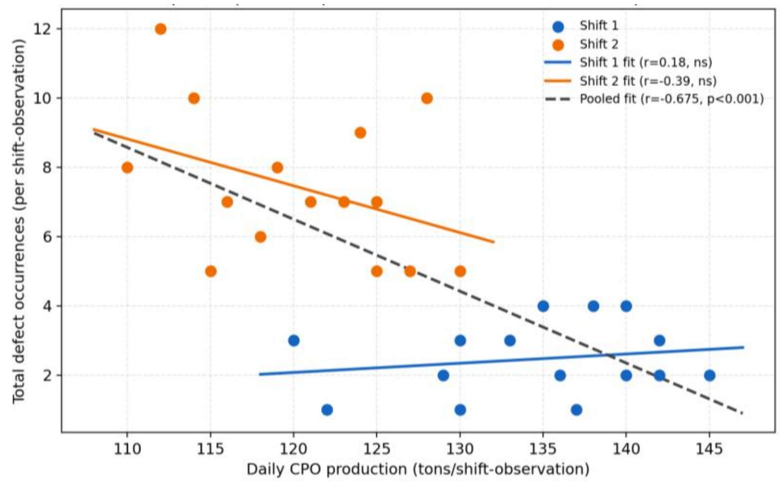

Figure 3. Production vs. Defect Occurrences, Stratified by Shift, with per-shift and pooled trend lines

The totals 2,019 + 1,807 = 3,826 tons are internally consistent. With 15 observations per shift, the averages are approximately 134.60 and 120.47 tons per shift-observation.

Process Flow and Fishbone Analysis

The quality-control flow links raw-material acceptance, sterilization, extraction, clarification, laboratory testing, and release or storage decisions. Quality deviations can occur before laboratory testing; therefore, QC should function as process control rather than only as end-product inspection.

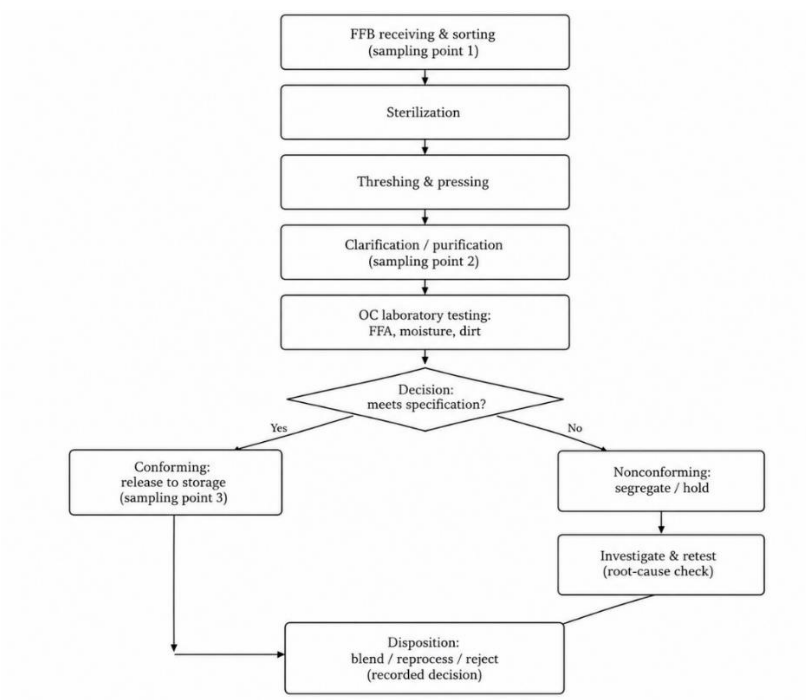

Figure 4. CPO Production and Quality-Control Flow

Potential causes are grouped into material, method, machine, human, and environmental factors. Material factors include immature or improperly processed FFB; method factors include sterilization and process settings; machine factors include press and clarification equipment; human factors include SOP compliance and sampling; environmental factors include cleanliness [20], [21].

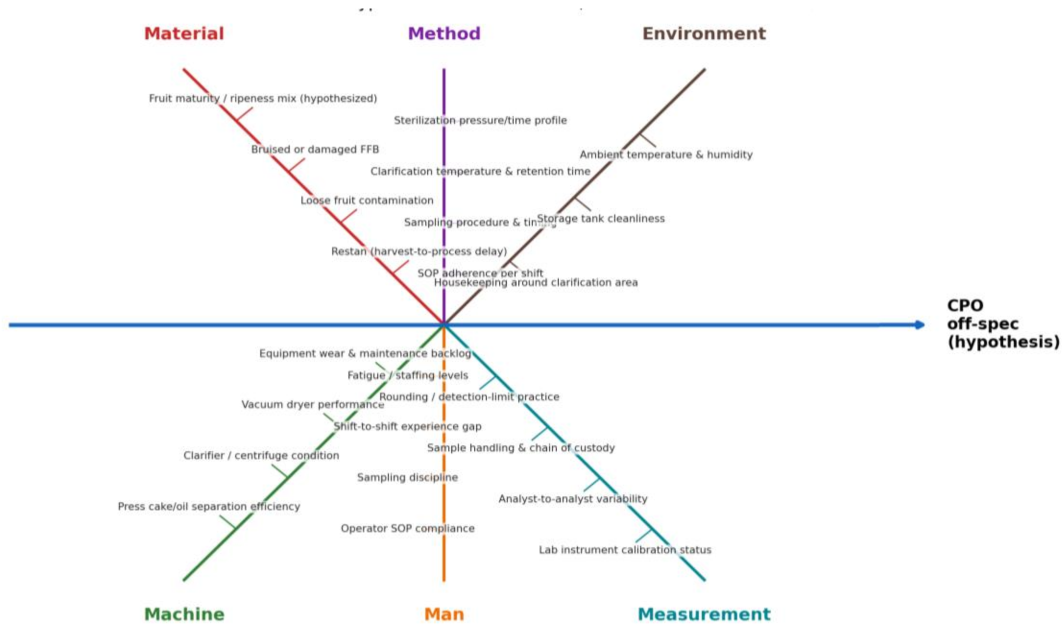

Figure 5. Fishbone (6M) Diagram of Hypothesized Causes of CPO Off-Spec Occurrences

Consistent with the labeling in Figure 5, every item in the fishbone diagram is reported here as a candidate cause identified through team discussion, not as a verified root cause. The material available for this study does not report who participated in generating these candidate causes, what production, maintenance, or laboratory records were reviewed, or whether any candidate cause was checked through Gemba (direct process) observation, the 5 Why technique, or quantitative process data; this should be reported by the author, and each recommended action in Discussion is explicitly tied back to the specific candidate cause it addresses, pending that verification.

DISCUSSION

Rational Subgrouping and the Corrected Special-Cause Finding

The principal statistical issue in the original analysis is that two systematically different strata, Shift 1 (rate 4.93%) and Shift 2 (rate 14.80%), were pooled into a single p-chart before assessing whether pooling was appropriate. Standard SPC guidance treats rational subgrouping, choosing subgroups likely to be internally homogeneous, as a precondition for a valid control chart [13], [14]; pooling two shifts with visibly different rates violates this precondition and can manufacture apparent instability that a correctly stratified analysis does not support. As shown in Results, the single point that exceeded the pooled UCL (24.00%) falls well within Shift 2's own UCL (29.87%) once shift is treated as the subgrouping variable. The defensible statistical conclusion is that this January dataset does not, on the evidence available, contain a special-cause signal in either shift; the originally reported "18 January special cause" finding is withdrawn[22], [23].

Simpson's Paradox in the Production-Defect Relationship

The pooled correlation (r = -0.675) reflects a between-shift pattern, not a within-shift relationship: Shift 1 happened to both produce more and record fewer defects, and Shift 2 happened to both produce less and record more defects, which alone is sufficient to generate a strong negative pooled correlation even absent any real effect of production volume on defects. The near-zero, non-significant within-shift correlations (0.18 and -0.39) and the regression result, in which shift, not production, is the significant predictor, support this reading. The original manuscript's implicit suggestion that higher production is associated with fewer defects should not be carried forward; if anything, the shift-adjusted result reinforces that shift-level operating conditions, not throughput, are the more plausible driver of the pattern in this dataset, pending direct investigation of what actually differs between the two shifts[24], [25].

What Actually Differs Between Shifts?

Because equal sample counts (750 per shift) demonstrate consistent sampling intensity rather than comparable process conditions, and because adjacent shifts may share common fresh-fruit-bunch supply, carry-over equipment condition, and laboratory systems, this study cannot determine whether Shift 2's higher rate reflects different FFB condition delivered to that shift, start-up or shutdown operations, equipment or process-setting differences, sampling-practice differences, different laboratory personnel, or some combination. Before recommending operator training specifically for Shift 2, each of these alternative explanations should be checked against production and maintenance logs, FFB grading records, and laboratory assignment records; the fishbone diagram (Figure 5) lists these as parallel candidate causes precisely because the current data cannot adjudicate among them [26], [27].

FFA Pathways

FFA's dominant share of recorded occurrences (Table 1, Figure 1) is consistent with the general literature on FFA formation via lipase activity following fruit bruising and delayed sterilization [1], [6], [7], but this study did not measure fruit maturity, bruising, loose-fruit handling, transport delay, restan (harvest-to-processing) time, or sterilization pressure/time profiles directly. Naming immature or over-ripe fruit as the principal cause of the FFA occurrences recorded here would therefore go beyond the evidence collected; it is retained in Figure 5 as a candidate cause requiring verification against measured restan time and sterilization records, not as an established explanation.

Moisture and Dirt Pathways

Moisture deviations are most plausibly linked to clarification temperature, dilution water management, settling and centrifuge (purification) performance, vacuum-drying effectiveness, tank condensation, and storage-management practice; dirt deviations are most plausibly linked to clarification and purification separation efficiency, separator condition, sludge handling, and contamination, including contamination introduced during sampling itself [28], [29], [30]. Because the moisture (0.15%) and dirt (0.02%) limits used at this mill are stricter than commonly cited SNI figures (Introduction), sampling representativeness and laboratory measurement uncertainty near these narrow limits are also plausible contributors to some recorded occurrences and should be examined as an alternative to a purely process-based explanation, particularly for samples recorded only marginally above the limit [31], [32].

Severity, Not Only Frequency

Because the dataset records pass/fail occurrences rather than continuous FFA, moisture, and dirt values, a sample recorded marginally above a limit is currently treated identically to a sample far above it. This conceals the magnitude of deviation, which matters for prioritizing corrective action, understanding blending or reprocessing feasibility, and assessing customer or contractual risk; continuous laboratory values, if retained in the company's records, should be analyzed directly [33], [34].

Comparison with Comparable CPO Mills

Previous CPO Seven Tools and SQC studies report FFA, moisture, and dirt as central quality parameters, with the dominant parameter varying by plant [3], [4], [7], [8], [9], [10], [11]; the finding that FFA dominates recorded occurrences at this mill is therefore plausible and broadly consistent with this literature, while remaining plant-specific. None of these comparison studies, to the author's knowledge, tests whether shift or another subgroup confounds a reported production-defect or time-series relationship in the way demonstrated here; this is offered as this study's most transferable methodological contribution to that literature, independent of whether the specific numeric rates at this mill match other sites. Direct numeric comparison with other mills' occurrence rates is not attempted here because sampling design, specification limits (including the stricter internal moisture and dirt limits noted in Introduction), and reporting conventions differ across the cited studies and are not uniformly documented [20], [21].

Operational and Economic Considerations

Where company evidence permits, off-spec CPO's practical consequences, price deductions, blending with conforming stock, reprocessing, segregation pending retest, rejection risk, and any yield-quality trade-off from tighter FFB sorting or shorter holding time, should be quantified and weighed against the cost of the proposed corrective actions. Tighter sorting and shorter restan targets are not costless: they can reduce throughput, increase handling labor, and create scheduling pressure at the loading ramp; this study does not have the operational or financial data to evaluate this trade-off and recommends it as a specific input to any final improvement decision, alongside the same feasibility question for maintenance scheduling, staffing for peak-only interventions, and calibration frequency [35].

Limitations

1. This is a single-mill study covering 15 calendar days (30 shift-observations) in one month; findings should not be generalized to other mills, months, or the wider palm-oil industry.
2. The data hierarchy (15 days x 2 shifts x 50 samples) was reconstructed from internal consistency in the charts, not confirmed directly from raw records; the author should publish the underlying anonymized subgroup table.
3. Whether the 148 recorded occurrences correspond to 148 unique defective samples or fewer, overlapping samples is not established; the reported 9.87% is an occurrence rate, not a confirmed unique-defective-sample rate.
4. The pooled p-chart and pooled correlation are both affected by shift confounding, as shown above; the stratified reanalysis in this revision is itself only a Phase I (retrospective, trial) analysis of the same 15 days per shift, not a validated prospective monitoring baseline.
5. The dataset consists of pass/fail counts rather than continuous FFA, moisture, and dirt measurements, so it cannot show process centering, dispersion, distance from specification, or support a capability index.

6. Neither the Poisson/exponential-type assumptions implicit in count-based control charting nor the specific distributional behavior of the underlying continuous parameters were tested, since continuous values are not available.
7. Fishbone causes are unverified hypotheses: no Gemba observation, 5 Why analysis, or process-record review is documented for any candidate cause in the material available.
8. Sampling procedure, laboratory method, measurement uncertainty, and calibration/traceability practice are not documented, which limits confidence in occurrences recorded close to the (comparatively strict) moisture and dirt limits.
9. No cost, staffing, land, or operational feasibility analysis accompanies the improvement recommendations below; they are prioritized by hypothesized impact, not by verified cost-benefit analysis.

Improvement Recommendations (5W+1H)

The following priorities are framed as evidence-bounded hypotheses for verification and pilot testing, not as confirmed corrective actions, consistent with the limitations above. Each should be evaluated prospectively (for example, through a before-after comparison or a structured DMAIC cycle) with new Phase II data rather than assumed to be effective based on this study alone.

Priority	What	Why (hypothesis)	Where/When	Who	How
1	Verify FFA pathway and tighten FFB sorting	Reduce high-FFA raw material if restan time and bruising are confirmed contributors	Loading ramp; continuous, pending verification	QC + production	Measure restan time and grading first; then strengthen maturity grading and target restan under 24 hours (definition and basis to be documented by the author)
2	Investigate shift difference before acting	Identify which factor (FFB, equipment, settings, sampling, lab) actually differs by shift	QC review meeting; before further shift-specific action	QC manager + shift supervisors	Compare shift-level FFB grading, equipment logs, and lab assignment records
3	Risk-based maintenance and calibration	Reduce moisture and dirt carry-over if linked to equipment condition	Pressing & clarification stations; frequency based on risk, not a fixed monthly default	Maintenance + QC	Inspect wear parts, clean separators, verify instrument calibration against manufacturer guidance and failure history
4	Standardize and verify sterilization profile	Reduce FFA-formation risk if sterilization time/pressure is shown to vary	Sterilizer; continuous, pending verification	Process supervisor	Log and review pressure/time profile against specification; calibrate instruments
5	Targeted QC and sampling training	Improve SOP and sampling compliance if training gaps are confirmed, not assumed	Mill site; periodic	HRD + Mill Manager	Refresh sampling SOP, critical limits, recording and feedback; confirm gap exists before training is delivered as the primary fix
6	Structured review of any future flagged shift-observation	Prevent recurrence and build a genuine Phase II baseline	QC meeting; triggered by a verified out-of-control point on the stratified chart	QC + production	Review shift, FFB, machine, process, and laboratory records for the specific flagged observation

CONCLUSION

This conclusion is limited to PT Tandan Sawit Papua, Arso Timur, and the verified 15-day (30 shift-observation) January 2026 period studied; it should not be generalized to other mills, months, or the palm-oil industry generally.
Seven Tools analysis of this dataset recorded 148 parameter-level off-spec occurrences against 1,500 tested samples, a 9.87% occurrence rate; whether this corresponds to 9.87% of samples being uniquely defective depends on unverified overlap among the three parameters and should not yet be stated as a unique-defective-sample rate. FFA/ALB was the largest recorded category (92 occurrences, 62.16%), followed by moisture (37; 25.00%) and dirt (19; 12.84%); this ranks where occurrences were recorded and should not, on its own, be read as identifying the underlying cause.

Shift 2 recorded substantially more occurrences than Shift 1 (111 versus 37) on lower production volume; this is reported as an operational signal requiring investigation of fresh-fruit-bunch condition, equipment, process settings, sampling practice, and laboratory conditions by shift, not as evidence of weaker operator discipline in Shift 2. After reconstructing the dataset's true day/shift structure and recomputing shift-specific control limits, the originally reported special-cause finding for "18 January" (24.00%, exceeding the pooled UCL of 22.52%) does not hold: that observation belongs to Shift 2's ninth day and falls within Shift 2's own UCL of 29.87%. Similarly, the pooled correlation between production and defects (r = -0.675) reflects confounding by shift (Simpson's paradox) rather than a genuine production-defect relationship; within-shift correlations are small and not statistically significant, and a shift-adjusted regression identifies shift, not production volume, as the significant predictor of defect counts.

The fishbone actions proposed (material, method, machine, man, environment, and measurement) are hypotheses to be verified through process records, root-cause analysis, and pilot corrective actions, not confirmed causes. Priority next steps are to collect continuous laboratory values in place of pass/fail counts, document fruit-condition and restan (harvest-to-processing delay) data, sterilization and clarification process variables, equipment maintenance records, and shift-level operating conditions, and to follow up with a properly stratified, prospective (Phase II) monitoring design once a stable baseline is established, before further capital investment or company-wide corrective action is finalized.

ACKNOWLEDGMENT

The authors thank PT Tandan Sawit Papua, Arso Timur, for providing the production and quality-control data used in this study, and acknowledge the support of the Industrial Engineering Program, Faculty of Science and Technology, Universitas Internasional Papua.

REFERENCES

[1] B. Herwanto and M. Effendi, “Quality Analysis of Crude Palm Oil (CPO) Processing and Proposed System Improvements to Increase Free Fatty Acid (FFA) Quality at PT. TLDN (CAP),” *J. IPTEK*, vol. 30, no. 1, 2026.

[2] S. Emebu, “Influence of moisture content, temperature, and time on free fatty acid in stored crude palm oil,” *Sci. Rep.*, vol. 12, p. 9846, 2022, doi: 10.1038/s41598-022-13998-1.

[3] R. Z. Syafira, S. H. Anwar, and Z. F. Rozali, “Pengendalian Mutu Crude Palm Oil (CPO) Dengan Metode Control Chart dan Failure Mode and Effect Analysis (FMEA) Pada Pabrik Kelapa Sawit PT. XYZ,” *J. Teknol. dan Ind. Pertan. Indones.*, vol. 14, no. 2, pp. 55–64, 2022.

[4] D. Kurniawan and W. Setiafindari, “Analisis Kualitas Produk Crude Palm Oil Menggunakan Metode Seven Tools di PT DK,” *Widya Tek.*, vol. 21, no. 2, pp. 88–96, 2022, doi: 10.33508/wt.v21i2.4525.

[5] D. Chandrahadinata and W. Nurdiana, “Analisis Pengendalian Kualitas Pada Crude Palm Oil untuk Meningkatkan Kualitas di PT. Condong Garut,” *J. Kalibr.*, vol. 19, no. 1, pp. 1–9, 2022, doi: 10.33364/kalibrasi/v.19-1.1045.

[6] E. Edyson, “Preprocessing Factors Affected Free Fatty Acid Content in Crude Palm Oil Quality,” *J. Ilmu Pertan. Indones.*, vol. 27, no. 2, pp. 177–181, 2022, doi: 10.18343/jipi.27.2.177.

[7] “Free Fatty Acid Formation Points in Palm Oil Processing and the Impact on Oil Quality,” *Agriculture*, vol. 13, no. 5, p. 957, 2023, doi: 10.3390/agriculture13050957.

[8] M. Hasanah and H. Haswan, “Analisis Pengendalian Mutu Crude Palm Oil (CPO) dengan Parameter Free Fatty Acid (FFA), Moisture dan Dirt Menggunakan Metode Statistical Quality Control (SQC) di PT Sari Buah Sawit Pasaman Barat,” Politeknik ATI Padang, Padang, Indonesia, 2023.

[9] J. D. Jaya and Khairiah, “Quality Evaluation and Causal Analysis of Crude Palm Oil Based on Critical Quality Parameters: Free Fatty Acids, Moisture, and Impurities,” *J. Teknol. Agro-Industri*, vol. 12, no. 1, pp. 1–10, 2025, doi: 10.34128/jtai.v12i1.227.

[10] A. L. Safitri and M. Husnah, “Analisis Kualitas Crude Palm Oil (CPO) Berdasarkan Parameter Free Fatty Acid (FFA), Moisture dan Dirt,” *Pascal J. Phys. Sci. Learn.*, vol. 9, no. 2, pp. 77–85, 2025.

[11] N. Apriani, S. R. Sibuea, and A. Hasibuan, “Pengendalian Mutu Crude Palm Oil (CPO) Menggunakan Metode Seven Tools pada PT. Perkebunan Nusantara IV Ajamu,” *Bul. Utama Tek.*.

[12] H. M. Siregar, A. Ishak, and H. A. Sinaga, “Quality control of crude palm oil (CPO) using DMAIC and fuzzy failure mode and effect analysis,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 801, p. 12121, 2020, doi: 10.1088/1757-899X/801/1/012121.

[13] D. C. Montgomery, *Introduction to Statistical Quality Control*, 8th ed. Hoboken, NJ: Wiley, 2020.

[14] J. Oakland, *Statistical Process Control*, 7th ed. Abingdon, UK: Routledge, 2020.

[15] M. J. Antony, *Lean Six Sigma for Small and Medium Sized Enterprises*, 2nd ed. Boca Raton, FL: CRC Press, 2020.

[16] D. Besterfield, *Quality Control*, 9th ed. Boston, MA: Pearson, 2020.

[17] P. Rogeboz *et al.*, “Oil Adulteration Evaluation Using High Performance Thin Layer Chromatography,” *Food Anal. Methods*, vol. 17, pp. 1336–1347, 2024, doi: 10.1007/s12161-024-02659-w.

[18] “Qualitative parameters and deterioration kinetics of palm oil, shea butter and their blend use for frying

cheese," *BMC Biotechnol.*, vol. 25, 2025, doi: 10.1186/s12896-025-01068-z.

[19] P. Laddawong *et al.*, "Federated Learning for Grading Oil Palm Fruit Ripeness in the Oil Palm Industry," *Appl. Comput. Intell. Soft Comput.*, 2025, doi: 10.1155/acis/9241523.

[20] Ng and V. Baeten, "Refined, bleached, deodorized palm oil adulteration screening using near infrared red spectroscopy," *J. Sci. Food Agric.*, 2025, doi: 10.1002/jsfa.14435.

[21] G. B. R. Bomfim, W. A. dos Santos, L. S. Sousa, P. P. L. G. Tavares, and M. E. de O. Mamede, "Physicochemical and sensory evaluation of the quality of crude palm oil (Elaeis guineensis) produced artisanally and industrially in Brazil," *Semin. Ciências Agrárias*, vol. 45, no. 5, pp. 1505–1520, 2024, doi: 10.5433/1679-0359.2024v45n5p1505.

[22] E. Z. H. Ooi *et al.*, "Underutilised palm stearin as hard stock for deep-frying medium and its performance for oil uptake in instant noodles," *J. Sci. Food Agric.*, 2024, doi: 10.1002/jsfa.13278.

[23] How and W. X. Soo, "Effect of butylated hydroxyanisole and α-tocopherol on the oxidative stability of palm olein during microwave heating," *Int. J. Food Sci. Technol.*, 2024, doi: 10.1111/ijfs.17374.

[24] A. M. Alhaji, E. S. Almeida, C. R. Carneiro, C. A. S. da Silva, S. Monteiro, and J. S. R. Coimbra, "Palm Oil (Elaeis guineensis): A Journey through Sustainability, Processing, and Utilization," *Foods*, vol. 13, no. 17, p. 2814, 2024, doi: 10.3390/foods13172814.

[25] N. Ollinger *et al.*, "High-Oleic Sunflower Oil as a Potential Substitute for Palm Oil in Sugar Coatings—A Comparative Quality Determination Using Multispectral Imaging and an Electronic Nose," *Foods*, vol. 13, no. 11, p. 1693, 2024, doi: 10.3390/foods13111693.

[26] D. C. Montgomery, *Design and Analysis of Experiments*, 10th ed. Hoboken, NJ: Wiley, 2020.

[27] S. George, *The Lean Six Sigma Pocket Toolbook*. New York, NY: McGraw-Hill, 2021.

[28] A. Pramana and E. Rossi, "Using palm oil residue for food nutrition and quality: from palm fatty acid distillate to vitamin E toward sustainability," *J. Sci. Food Agric.*, vol. 105, no. 9, pp. 4728–4740, 2025, doi: 10.1002/jsfa.13878.

[29] A. Hayyan, K. M. Abed, M. Hayyan, A. J. Ali, W. J. Basirun, and S. S. Syed Putra, "Direct application of tungstosilicic acid hydrate for the treatment of high free fatty acid in acidic crude palm oil and for biodiesel production," *J. Am. Oil Chem. Soc.*, vol. 102, no. 2, pp. 375–384, 2025, doi: 10.1002/aocs.12896.

[30] W. N. A. W. Osman, J. K. L. Min, and S. Samsuri, "Analysis of chemical properties for biodiesel derived from crude palm oil," *Can. J. Chem. Eng.*, vol. 103, no. 5, p. 2053, 2025, doi: 10.1002/cjce.25501.

[31] I. O. for Standardization, "ISO 9001:2015 Quality Management Systems - Requirements," 2015, *ISO, Geneva, Switzerland*.

[32] B. S. N. (BSN), "SNI 01-2901-2006: Minyak Kelapa Sawit Mentah (Crude Palm Oil)," 2006, *BSN, Jakarta, Indonesia*.

[33] F. A. binti Marzuki, M. Rahman, and A. Manaf, "Effect of aqueous washing on chloride reduction and the storage stability on low-quality crude palm oil prior to refining," *Food Addit. Contam. Part A*, vol. 42, no. 5, pp. 580–593, 2025, doi: 10.1080/19440049.2025.2484595.

[34] A. Hasliyanti, C. M. Che Rahmat, and M. H. Ng, "Determination of potential chloride contaminants throughout palm oil milling by 35-Cl nuclear magnetic resonance," *J. Food Compos. Anal.*, vol. 141, p. 107368, 2025, doi: 10.1016/j.jfca.2025.107368.

[35] "Comprehensive insights into the detoxification effects of refining process on hazardous PAHs in 13 crude vegetable oils," *J. Food Compos. Anal.*, vol. 145, p. 107809, 2025, doi: 10.1016/j.jfca.2025.107809.