ISSN: 2754-6659

Journal of Artificial Intelligence & Cloud Computing

SCIENTIFIC Research and Community

Review Article

Open Access

From Black-Box Prediction to Conservation-Aware Decisions: Physics-Informed Ai for Palm Oil Mill Effluent and Circular Biorefineries

Loso Judijanto

IPOSS Jakarta, Indonesia

ABSTRACT

Palm oil mill effluent (POME) is both a high-strength wastewater-management challenge and a prospective feedstock for circular biorefineries. Although machine-learning models increasingly predict methane production, effluent quality, and process performance, accuracy on historical observations does not ensure compliance with chemical balances or safe behavior during disturbances. This qualitative literature review critically synthesizes research from 2020–2026 on industrial POME digestion, physics-informed and hybrid machine learning, anaerobic biokinetics, soft sensing, advanced process control, and multi-product resource recovery. It distinguishes published industrial-data demonstrations from proposed plant-wide decision architectures. The synthesis identifies six themes: incomplete process observability, the conditional value of mechanistic constraints, uncertain generalization across mills, opportunities for constrained supervisory control, cross-stream resource interactions, and governance requirements for deployable AI. A Palm Circular Biorefinery Digital Intelligence framework is proposed, combining auditable measurement, substrate-specific mechanistic models, data-driven residual learning, uncertainty estimation, resource-balance accounting, and independently enforced operational limits. The review argues that conservation-aware learning should be evaluated against strong mechanistic and conventional machine-learning baselines using temporal and cross-site tests, constraint violations, calibration error, operating costs, and verified environmental outcomes. It prioritizes shadow-mode and human-authorized trials before bounded automation. The proposed framework is a research and implementation agenda rather than evidence of an already validated autonomous palm biorefinery.

***Corresponding author**

Loso Judijanto, IPOSS Jakarta, Indonesia

Received: October 30, 2026; **Accepted:** November 03, 2026; **Published:** November 13, 2026

Keywords: Physics-Informed Artificial Intelligence, Palm Oil Mill Effluent, Anaerobic Digestion, Hybrid Modeling, Biogas; Methane, Soft Sensors, Model Predictive Control, Biochar, Circular Biorefinery

JEL Classification: C45; C61; L66; Q25; Q53.

Introduction

Background: Industrial Residues as Interconnected Process Resources

Palm oil mills operate at the boundary between agricultural biomass supply, thermal and mechanical manufacturing, biological treatment, and environmental compliance. The liquid waste stream known as palm oil mill effluent (POME) contains biodegradable organic matter, residual oil, suspended solids, and nutrients, but its composition varies with extraction operations, dilution, and upstream handling. A conventional treatment perspective regards these flows primarily as discharge liabilities. A circular-biorefinery perspective instead treats the same flows as potential sources of methane, recoverable water, nutrients, residual oil, and carbon-rich materials, while recognizing that recovery equipment generates additional resource demands and residuals. Neither framing by itself guarantees environmental benefit: avoided emissions and recovered products must be established through a transparent boundary and credible comparison with the displaced alternative. Recent POME treatment and valorization studies identify these competing opportunities and operational constraints, rather than a single universally superior recovery pathway [1-4].

The rapid expansion of process analytics creates an attractive but risky shortcut. A model trained on historical operating data may report excellent methane-yield prediction while failing when a digester experiences unusual loading, a shift in feedstock composition, chemical inhibition, or faulty sensors. Industrial POME studies have applied neural networks, response-surface models, tree methods, and data-assisted simulation to predict biogas production and optimize digestion or co-digestion. These investigations demonstrate substantive progress, but their published results relate to specific operating data, model choices, and validation settings. Prediction, retrospective optimization, and industrial closed-loop control are distinct evidence categories. The relevant research problem is therefore not merely whether artificial intelligence can fit biogas data, but whether a process-aware learning system can generate reliable recommendations in the presence of biochemical constraints, changing operating conditions, and incomplete observations [5-9].

Urgency and the Case for Physics-Informed Learning

Physics-informed artificial intelligence offers one response to the fragility of purely empirical optimization. Its broad family includes models that penalize violations of known differential equations; serial hybrids that estimate unknown kinetic parameters and feed them into first-principles simulators; parallel models that correct systematic mechanistic errors; and constrained surrogates

designed to preserve material, energy, or elemental balances. Their usefulness is conditional rather than automatic. If a mechanistic equation is poorly specified for a particular substrate or wastewater facility, forcing data to comply with it may increase bias; if enough representative operating data are available, a carefully regularized conventional model may already perform well. The question is how much trustworthy physics should enter which part of the model, and at what computational and organizational cost [10-14].

The POME case is unusually instructive because a directly relevant industrial-data study integrated reduced anaerobic-digestion equations with neural learning. Shaw and colleagues modeled three outputs at two industrial digestion plants and found that their physics-informed approach improved reported test accuracy and training efficiency relative to specified alternatives. Their result supports the plausibility of model fusion, not the conclusion that every digesting wastewater stream needs an identical network or that online control has been experimentally validated. Complementary research has explored machine-learning-guided co-digestion with decanter cake, biochar additions, and biological scrubbing for hydrogen sulfide removal. The diversity of these cases makes POME an informative platform for analyzing model form, operating risk, resource synergies, and the gulf between strong retrospective predictions and dependable decisions [15-18].

Research Gap, Questions, Objectives, and Contribution

The first gap is architectural: empirical POME predictors are often scoped to one output or unit operation, whereas a circular biorefinery requires coordination among liquid-effluent treatment, biogas conditioning, power generation, solid-residue conversion, and water or nutrient recovery. The second gap is epistemic: many studies measure predictive fit without reporting whether estimates preserve physical conservation, remain calibrated under a new mill's conditions, or identify model inadequacy. The third gap is institutional: a proposed supervisory optimizer must operate alongside certified plant control, environmental permits, equipment limitations, and operators with legitimate authority to reject unsafe actions. The article therefore assesses the promise and limits of conservation-aware modeling instead of treating digitalization as a sufficient condition for circularity [19-22].

Five research questions guide the synthesis. RQ1 asks which mechanistic assumptions and conservation relationships are defensible for data-assisted modeling of POME anaerobic digestion and biogas treatment. RQ2 asks under which evidence and data conditions physics-informed or hybrid models can offer value beyond conventional data-driven and mechanistic baselines. RQ3 asks how soft sensors, uncertainty estimates, and sequential decision methods can be converted into safe supervisory recommendations. RQ4 asks how POME-derived biogas, decanter cake, empty fruit bunches, biochar, water, and energy can be represented as coupled material and environmental flows. RQ5 asks what validation, governance, and economic tests are required before a cross-mill digital-intelligence platform can be considered operationally credible. The objective is to answer these questions through critical thematic interpretation and to formulate a falsifiable Palm Circular Biorefinery Digital Intelligence framework [19,23,24].

Literature Review: Conceptual and Theoretical Foundations

Mechanistic Anaerobic Digestion and The Limits of First Principles

Anaerobic digestion is not a single reaction that converts chemical oxygen demand into methane at a fixed rate. Hydrolysis, acidogenesis, acetogenesis, and methanogenesis involve interacting microbial populations, substrate fractions, intermediate products, inhibition mechanisms, gas-liquid transfer, and reactor hydraulics. Mechanistic models represent selected states and rate expressions and thereby make assumptions inspectable. They may be valuable for testing whether an input-output response is biologically plausible, but calibrated parameters can be non-identifiable when important microbial states or feed characteristics are unmeasured. Simplified models additionally compress microbial heterogeneity into aggregate reactions. A framework for industrial POME must balance mechanistic completeness against observability, parameter-identification burden, and real-time computational demands [22,23,25,26].

This trade-off becomes acute for variable feedstocks. Changing organic loading rate alters not just expected gas output but also the timing of intermediate accumulation, buffering demand, and risk of process souring. Hydraulic retention time interacts with reactor design, active biomass retention, and substrate biodegradability; a control policy that recommends a very short residence time solely because instantaneous gas output is high may compromise COD removal or long-run stability. Industrial lagoon or covered-digester operations have further complications: unmeasured mixing heterogeneity, rain dilution, temperature variation, and partially known sludge inventories. A first-principles model is therefore a structured approximation that requires site-specific testing rather than an authoritative oracle [6,15,25,27].

Black-Box, Gray-Box, and Physics-Informed AI

Black-box predictive models learn input-output mappings from available examples. Tree ensembles can tolerate heterogeneous numerical variables and sometimes work well on medium-sized plant datasets; neural networks can approximate nonlinear effects when training information is adequate. Their weakness is not the absence of utility but the absence of automatic physical guarantees. A network may predict negative concentrations, exceed a theoretically available product balance, or extrapolate unrealistically beyond observed feed rates unless its architecture or outputs are restricted. By contrast, hybrid or gray-box modeling makes an explicit division of labor between physical representation and learned components. One may learn difficult reaction-rate corrections, infer missing states, estimate site-specific parameters, or correct residual prediction errors while conserving the quantities that are well established [9,11,28,29].

Physics-informed neural networks are a narrower subset of this broader scientific machine-learning landscape. A differential-equation residual may contribute a weighted term to the training objective, but a small average residual does not alone establish reliable initial-condition behavior, conservation over long trajectories, or safe extrapolation. Physics-informed process modeling also encounters stiff chemical and biological dynamics, gradients that compete during training, uncertain parameters, and missing equations for poorly characterized states. Hybrid frameworks differ substantially in how physics enters a model; an operational review must compare them on equal predictive tasks and identify whether the imposed rule is an exact identity, a tested approximation, or a provisional hypothesis [10,12,13,30,31].

Neural Differential Equations, Surrogates, and Mechanistic Residual Learning

Neural ordinary differential equations learn continuous-time state derivatives or corrections and are attractive when measurements arrive irregularly, an issue familiar to industrial wastewater plants. A neural component can describe otherwise unmodeled microbial effects while a process solver enforces known accumulation balances. Nevertheless, solver stiffness, poorly constrained

latent states, and error accumulation can make unconstrained neural differential equations difficult to interpret. Process-control studies have investigated how such models can be embedded within model-predictive control, although application to a new POME installation requires independent state-observability and computational-speed demonstrations. Reactor-modeling work also suggests that conservation-regularized learning may improve robustness in certain noisy or data-scarce conditions [29,32-35].

Surrogate models should be assessed in relation to their intended role. A fast emulator may replace repeated runs of a slow first-principles model during scenario screening without being granted authority to issue actuator commands. A soft sensor may estimate COD, volatile fatty acid trends, or other laboratory-intensive variables between physical assays, but it still requires a credible link between its predictors and the target quantity. A planning model that allocates biogas among electricity generation, boiler support, or upgrading faces a different uncertainty structure from a biological-state estimator. Treating all these models as interchangeable because they employ neural networks obscures the engineering requirements that determine trustworthiness [19,36-38].

Closed-Loop Process Control, Safety, and Explanatory Accountability

Advanced process control is best viewed as constrained decision-making over a moving horizon. A supervisory model proposes bounded changes, the plant carries out authorized actions, and subsequent observations reveal whether the expected response occurred. Model-predictive control has a strong conceptual fit because it can optimize projected trajectories subject to input and state limits; reinforcement learning is another option for repeated decisions but does not independently establish safety. Both depend on the quality of their state estimate, transition model, and feasible-action envelope. For POME digestion, candidate supervisory decisions could include gradual feed-rate changes, co-substrate blending, gas routing, or scrubbing adjustments, whereas emergency shutdown logic and certified equipment interlocks should remain outside a learned reward or objective [13,32,39,40].

Interpretation and uncertainty matter for operators. A prediction can be numerically accurate on a test set while being sensitive to an unnoticed sensor substitution or an out-of-distribution inflow. Reliable decision support should distinguish measured values from estimated states, specify key assumptions and binding constraints, report prediction intervals when justified, and abstain from prescriptive advice when uncertainty becomes excessive. Explainable models can show influential variables, but feature attribution is not a substitute for a causal demonstration of process benefit. Industrial wastewater and general digestion research support combining explainability with process expertise, rather than allowing importance rankings to substitute for chemistry [9,30,31,36,41].

Circular Biorefinery Theory: Mass Allocation and Consequential Accounting

Circularity depends on the allocation of real material, carbon, water, nutrient, and energy flows. POME digestion generates a gas stream containing methane and carbon dioxide, a residual liquid and solids stream, and treatment obligations that do not disappear when energy is recovered. Palm solid residues can be directed to combustion, composting, biochemical conversion, or pyrolysis, but the same ton of biomass cannot simultaneously support every route. Biochar used as a digester additive might improve selected biological conditions while consuming a residue otherwise available for other uses; producing and handling the biochar can also have environmental burdens. These trade-offs must be represented explicitly when AI chooses process configurations or feedstock allocations [2,3,17,21,42].

Water reclamation, membrane concentration, residual oil recovery, and biological polishing increase the possible products of a POME-centered biorefinery, yet they impose treatment quality, fouling, pumping, cleaning, concentrate-disposal, and energy constraints. Economic models should therefore distinguish gross biogas value from net realized value after purification, downtime, chemical dosing, maintenance, and financing. The environmental calculation should also distinguish on-site methane capture from net lifecycle emissions under a documented counterfactual. This systems perspective motivates multi-objective optimization in which water recovery and nutrient reuse are not presumed to be environmentally beneficial merely because they appear on a technology diagram [4,20,24,43-48].

Method: Qualitative Literature Review

Review Design and Epistemic Scope

This article uses a qualitative literature review (QLR), not a systematic literature review or meta-analysis. Its purpose is to integrate distinct lines of evidence, compare claims against the conditions under which they were generated, and develop a process-engineering argument that can guide subsequent measurement and experimentation. The synthesis engages with peer-reviewed research published from 2020 through October 2026, including POME process studies, hybrid-AI methodologies, digestion biology, digital twins, control engineering, and circular resource recovery. It does not present a PRISMA flow diagram, claim exhaustive database retrieval, assign numerical quality scores, or calculate a pooled intervention effect. This distinction avoids giving narrative interpretation an unsupported appearance of exhaustive screening [19,49,50].

The review gives greatest weight to evidence with clearly described physical systems and operational measurements, while also using methodological work from adjacent chemical, biochemical, and wastewater-processing domains where a transferable design principle is relevant. Adjacent-domain evidence is explicitly interpreted as methodological support, not proof that the same improvement holds at a palm oil mill. A positive result in a pharmaceutical bioreactor, a neural-ODE reactor example, or an urban wastewater facility helps identify plausible modeling strategies but cannot substitute for POME-specific hydrodynamics, feed chemistry, microbial behavior, and local operating conditions [13,32,38,51,52].

Literature Identification, Appraisal, and Thematic Coding

Source identification combined targeted journal and DOI searches with forward and backward exploration of core articles on physics-informed POME digestion, POME co-digestion, biological gas treatment, and hybrid bioprocess modeling. Sources were retained when they contributed a verifiable empirical result, a mechanistic model, a comparative method, or a clearly articulated implementation constraint relevant to the research questions. The appraisal recorded whether each contribution involved laboratory data, industrial historical data, independent site validation, actual intervention, or a conceptual review. Publication date and engineering relevance were necessary, but model complexity and reported accuracy were not treated as substitutes for external validity [7,15,18,22,23].

Three rounds of thematic analysis were applied. The first organized evidence by physical unit operation and recovered stream; the second distinguished data-only, mechanistic, hybrid, and conservation-aware approaches; and the third examined decision readiness through observability, constraints, uncertainty, baseline strength, and prospective verification. A claim about better out-of-sample prediction was separated from a claim about safer operation, and neither was accepted as evidence for environmental improvement without appropriate outcome measurement. These interpretive rules were used to derive the six findings below and the proposed framework. The resulting synthesis is conceptually accountable but does not claim statistical representativeness of all available publications [1,9,11,19].

Results: Six Thematic Findings

Theme I: POME predictability has advanced faster than online observability

Published industrial-scale analyses demonstrate that neural networks, ensemble methods, and response-surface models can capture important relationships between loading, pH, temperature, substrate blends, and biogas-related outputs. The results encourage data-assisted digestion management, particularly because industrial POME systems operate under greater feedstock variation than tightly controlled laboratory reactors. Yet laboratory assays and process historians do not sample every relevant microbial state or effluent characteristic. Training labels can be delayed, irregular, or indirectly inferred; missing observations can conceal transient instability; and equipment maintenance may change the statistical meaning of a sensor. An accurate static regression is therefore not equivalent to a continually synchronized process state [5-7,15,36].

A useful comparison distinguishes three levels of observational readiness. The first is recorded data with inconsistent timestamps and uncertain sampling protocols. The second is quality-controlled time series joined to batch, unit, and maintenance identities, with missingness flags and calibration information. The third adds validated soft sensors, probabilistic or interval state estimates, residual checks, and explicit detection of changing process regimes. The transition between these levels is operational rather than rhetorical. Independent wastewater digital-twin and soft-sensing studies indicate that physical instrumentation, process expertise, and data architecture remain central to successful analytics, even where machine-learning methods are technically sophisticated [19,36,37, 53,54].

Theme II: Mechanistic Information Helps Selectively, Not Universally

The most directly relevant demonstration is the physics-informed deep-learning analysis of two industrial POME digesters. By combining a reduced digestion model with neural learning, Shaw and colleagues reported stronger test performance and more efficient training than their specified mechanistic and neural comparators. This is persuasive component-level evidence that process equations can be productive regularizers for particular measured POME systems. It does not isolate a universal advantage for all physics-informed architectures, and it does not establish that a lower training loss corresponds to a stable controller under altered feedstock, unexpected inhibition, or a new plant's equipment configuration. The review accordingly treats physical knowledge as a testable design resource rather than a label conferring automatic superiority [10,13,14,15].

Other hybrid studies suggest how to choose the location of model fusion. Neural estimation of uncertain kinetic parameters can retain a familiar digestion balance, whereas a parallel residual model may correct an otherwise reasonable mechanistic forecast. Continuous-time neural models may be suitable for irregular samples but introduce additional stiffness and integration considerations. Research from bioprocess engineering emphasizes that more detailed physics can increase calibration demands and may not outperform a carefully chosen simpler model when the available observations do not identify the added states. A fair experiment must therefore compare black-box learning, calibrated mechanics, parameter-learning hybrids, residual hybrids, and constraint-aware networks under the same data and decision opportunities [11,12,23,29,38].

Theme III: Anaerobic Digestion is a Coupled Stability and Production Problem

Methane output is not an adequate isolated control target. A plant may generate more gas for a period while moving toward acid accumulation, foaming, biomass washout, or undesirable effluent concentrations. Methane fractions and H_2S conditions matter for engine operation, while high COD-removal performance can conflict with maximizing volumetric loading in certain regimes. Industrial co-digestion analyses show the potential importance of blending POME with decanter cake and jointly evaluating operating conditions, product outcomes, and economic performance. However, the significance of predicted gains depends on reactor design, feed properties, data representativeness, and treatment of operating costs, not simply the highest reported coefficient of determination [7,8,16,22].

Biochar-assisted digestion extends the problem from a single liquid feed to an allocation decision across liquid and solid residue chains. Two 2025 studies investigate how biochar dose and feed-to-inoculum or organic-loading conditions interact with methane performance; one reported a 104% methane-yield increase under its optimized experimental conditions. This percentage is an outcome of that study's setup, not a general improvement transferable to every industrial mill. A plant-scale recommendation would need to account for biochar origin, dose, reactor mixing, production cost, residual solids handling, and the opportunity cost of redirecting empty fruit bunches from other uses. Hybrid AI should therefore incorporate both digester kinetics and material-allocation accounting [17,24,27,42].

Theme IV: Gas Cleanup and Water Recovery Create Cross-Unit Dependencies

Gas generation does not directly equal usable energy. The biogas quality supplied to a generator, boiler, or upgrading unit depends on methane fraction, moisture, hydrogen sulfide, and other constituents, as well as the technical limits of downstream equipment. Industrial POME bioscrubber research combines process simulation with an artificial neural network and reports promising plant-specific predictions of H_2S removal. That work illustrates how observations in one unit can inform operating decisions in another, but it should not be interpreted as proof that a shared plant-wide controller has already been implemented or validated. An integrated model must distinguish digester gas production from recoverable gas after cleanup and compression or distribution losses [4,18,19].

Effluent polishing and water reclamation create comparable couplings. Membrane recovery can reduce fresh-water requirements in favorable configurations but may increase pumping energy, concentrate production, fouling, and chemical cleaning; recovering oil or nutrients changes the composition of streams returning to digestion or disposal. Algal systems introduce another possible link between carbon dioxide, light availability, nutrient uptake, and

product harvesting. These paths are not interchangeable, and not every mill has the same water demands or integration possibilities. The thematic result is that a useful optimization framework must represent unit-specific limitations and cross-stream consequences instead of treating each high-potential technology as an additive independent benefit [3,20,43-45].

Theme V: Prediction Accuracy Alone Cannot Establish Decision Quality

A model reporting a high R^2 value may still fail as an optimizer if errors cluster around abnormal loading, equipment failures, or the largest environmental risks. Temporal splitting is especially important in process histories, since neighboring timestamps share operating patterns, seasons, and control policies. Cross-mill testing is harder still: different wastewater composition, sampling practices, lagoons or reactors, retention behavior, and process control conventions can produce dataset shift. The correct benchmark must include temporal blocks, stressed operating periods, and ideally independent plants, with measurement of residual physical violations, confidence calibration, and computational latency in addition to point prediction [13,15,30,31,49,52].

Counterfactual optimization raises an additional limitation. Historical operators typically changed loading or dosing in response to observed disturbances; the available data therefore do not provide randomized evidence of what would have happened under an alternative policy. A trained surrogate can generate plausible scenarios, but its projected savings are not causal field effects without a validated transition representation and prospective evaluation. An apparent advantage may also disappear once a modern deterministic optimizer, calibrated mechanistic controller, or expert rule set is included as the comparator. The literature thus supports a hierarchy of evidence ranging from fit on historical records to independent-site prediction, shadow-mode recommendations, human-approved interventions, and finally bounded operational control [22,32,39,40,55].

Theme VI: Circular Outcomes and Governance Must Be Measured Together

A circularity claim cannot rest only on higher modeled methane yield. The relevant system may include avoided lagoon methane, displaced grid electricity or boiler fuel, digestion emissions, H_2S-removal inputs, biochar manufacturing, water treatment, residual solids, and by-product transport. Site-specific boundaries determine whether a policy reduces total emissions or merely moves burdens between operations. Studies of POME-derived energy and integrated mill operations make clear that process technology choices affect both environmental and economic accounting, even before machine learning is added. A proposed intelligence layer needs explicit material-balance, energy-balance, cost, and environmental ledgers, with measured values differentiated from uncertain modeled [2,4,21,24].

Governance extends the model boundary to technicians, environmental managers, and plant owners. Sensor drift, data tampering, informal changes to operating setpoints, or silent model updates can invalidate a prediction even when the underlying algorithm has not changed. The legitimate response is versioned data and models, audited recommendations, role-specific approvals, independent engineering interlocks, and the ability to reject or revert an action. A digital twin should reproduce observed physical discrepancies and suspend recommendations when its model no longer represents the process; an optimizer without such feedback is better described as offline decision support. The research literature provides strong reasons for these design choices, while field-validated demonstrations across a fully integrated POME circular biorefinery remain limited [19,37,40,56,57].

Discussion and Analysis

Interpreting The Value Proposition of Conservation-Aware Intelligence

The synthesis supports a conditional rather than categorical value proposition. Physics-informed intelligence is most defensible when the physical subsystem has reliable conservation relations, some partially understood dynamics, costly or sparse state measurements, and a decision problem whose consequences can be tested. In that setting, established balances can constrain the feasible region while machine learning approximates uncertain kinetic responses or unobserved effects. It is less compelling to add a complex neural architecture where a calibrated first-principles model or straightforward regression already performs comparably and makes fewer demands on maintenance personnel. The relevant measure of innovation is not the novelty of the optimizer but the quality, traceability, and operational usefulness of predictions under disturbed conditions [10,12,13,34].

One valuable distinction is between three types of physical knowledge. Exact accounting identities, such as conserved mass expressed in consistent units, should generally be enforced structurally or with very strong constraints. Empirical constitutive relations, such as approximate biodegradation response to temperature and substrate concentration, should be calibrated with uncertainty and revised when operating conditions shift. Operational rules, such as maximum allowable loading or H_2S concentration for an engine, are not simply model preferences; they belong in independently checked feasibility and safety layers. Confusing these categories can produce a network that reports excellent predictive accuracy while offering physically impermissible recommendations. This distinction also clarifies why 'physics-informed' is not a homogeneous technical standard or a guarantee of process safety [11,22,23,40,].

Modeling Reactor Behavior Without Obscuring Biological Uncertainty

For a well-mixed reactor, the dissolved-substrate inventory is represented conceptually by Equation (1). Here, V is reactor liquid volume, C is reactor substrate concentration, Qin and Qout are inlet and outlet volumetric flow rates, Cin and Cout are corresponding substrate concentrations, and rtotal is an effective volumetric substrate-consumption rate. The terms must be dimensionally consistent, and changes in volume must not be omitted if the reactor is not at constant volume [15,23].

$$\frac{\mathrm{d}(VC)}{\mathrm{d}t} = Q_{\mathrm{in}}C_{\mathrm{in}} - Q_{\mathrm{out}}C_{\mathrm{out}} - Vr_{\mathrm{total}} \qquad (1)$$

Equation (1) is a simplified accounting relationship rather than a complete anaerobic-digestion model. A process-specific model must add microbial intermediates, gas–liquid transfer, multiple substrate fractions, and appropriate sludge-retention dynamics. As shown in Equation (2), a hybrid representation may supplement calibrated mechanistic kinetics with a bounded data-learned residual, provided that its sign, dimensions, and implied material or elemental transfers remain physically defensible [10,23,35].

$$r_{\mathrm{total}} = r_{\mathrm{mech}}(\mathbf{x}, \boldsymbol{\theta}) + r_{\mathrm{res}}(\mathbf{x}, \mathbf{z}, \boldsymbol{\phi}) \qquad (2)$$

For multiple product streams, the same logic extends to carbon and chemical oxygen demand accounting over a specified period. Total carbon entering through POME, decanter cake, and other additions must be reconciled against carbon leaving in methane,

carbon dioxide, recovered products, residual effluent, solids, measured leaks, and inventory changes within the stated boundary. A practical constraint need not pretend that all measurements are exact: it can explicitly retain an uncertainty or unmeasured-loss term, require that its magnitude remain plausible, and trigger investigation when residuals trend beyond calibrated tolerances. This form of accountability is more useful than reporting a single accuracy score because it can expose sensor faults, omitted sources, or physically inconsistent model extrapolation [3,4,19,24].

Palm Circular Biorefinery Digital Intelligence Architecture

Figure 1 presents the proposed Palm Circular Biorefinery Digital Intelligence architecture. Its input layer joins feed and residue provenance, online sensor measurements, laboratory assays, equipment and maintenance records, and energy or water meters. A synchronized information contract attaches identity, time, measurement confidence, and plant context to those observations. The core modeling layer separates physical balances and calibrated anaerobic kinetics from learned residuals, estimated latent states, and uncertainty monitors. Application modules then examine digester stability, biogas quality, downstream cleanup, recovered water, and solid-residue utilization, while an external authorization layer limits how recommendations affect real equipment [15,19,36,46,56,58].

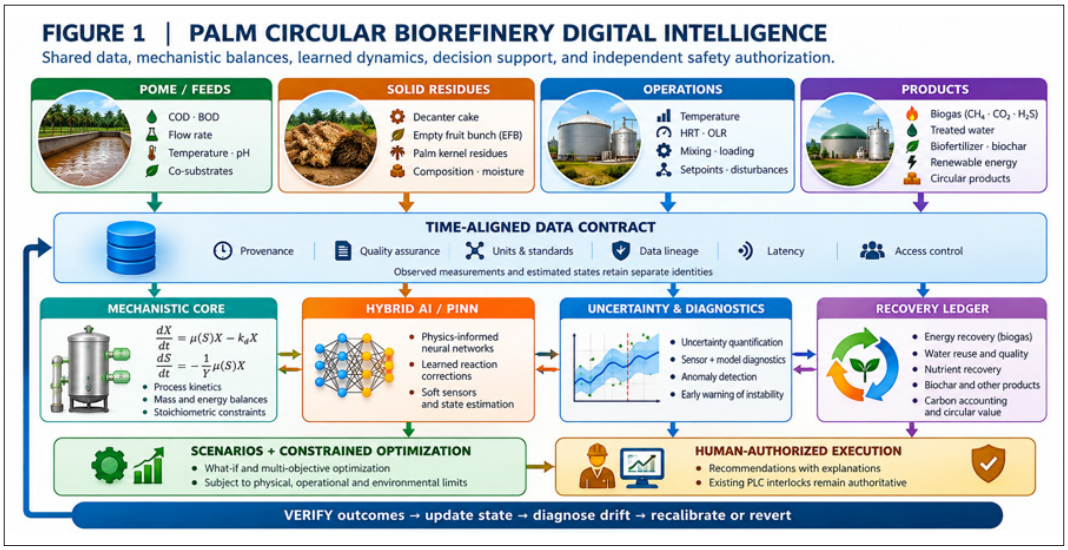


Figure 1: Proposed Palm Circular Biorefinery Digital Intelligence Architecture (Conceptual)

The essential architectural property is not universal automation but auditable coupling. The proposed physical models need not run at the same frequency: digesters may be monitored on a sub-hourly or daily basis, laboratory COD may be available less often, and solid-residue allocation may be revised weekly or seasonally. Data assimilation reconciles these clocks before any optimization uses them. When a methane prediction fails, the discrepancy must be attached to the corresponding feed batch, instrument status, and model version; when a gas-cleanup unit approaches a constraint, energy routing should update without falsifying the upstream fermentation state. Figure 1 therefore depicts an intended research architecture, not an already installed mill-wide control system [18,37,59,60].

Objective Design and Constraints for Circular Resource Allocation

A supervisory optimizer should assess net system value rather than a single intermediate product. Equation (3) expresses an illustrative monetary objective in which Renergy and Rmaterials denote net revenues from usable energy and recovered materials, and Coperations, Cenvironment, and Closses denote operating costs, monetized environmental burdens, and avoidable-loss costs, respectively. All terms must refer to the same accounting period and monetary units; nonmonetizable mandatory limits cannot be offset by revenue. The objective weights, accounting boundaries, and prices must be declared and tested through sensitivity analysis, while permitted operating ranges, environmental discharge standards, and equipment interlocks must be independently enforced [7,8,24,40].

$$J = R_{\text{energy}} + R_{\text{materials}} - C_{\text{operations}} - C_{\text{environment}} - C_{\text{losses}} \qquad (3)$$

Co-substrate selection introduces a portfolio problem. Increasing decanter-cake input can supplement organic material during low-effluent periods, but it also changes mixing, pretreatment energy, solids loading, and sludge disposal. Biochar addition affects both digester performance and the solid-residue processing chain; diverting empty fruit bunches to pyrolysis may reduce their availability as boiler fuel or soil amendment. Membrane polishing can deliver reusable water but may shift burdens to pumps, cleaning chemicals, or concentrate disposal. A multi-stream optimization should therefore calculate marginal value and residual impacts for each incremental allocation, explicitly preventing double counting of the same biomass or claimed emission reduction [16,17,20,27,42,43].

From Prediction and Soft Sensing to Safe Supervisory Optimization

Figure 2 shows how a prediction service could be embedded within a conservative control cycle. After the plant senses and time-aligns data, mechanistic and learned models construct candidate trajectories; a supervisory optimizer compares only feasible alternatives. A human decision-maker receives the expected gain, uncertainty bounds, physical-balance diagnostics, and reasons for any blocked actions. The existing plant automation executes authorized changes within certified process bounds, and post-action measurements are compared with the forecast. A large mismatch triggers investigation, recalibration, or a return to established operating rules rather than uncontrolled online retraining [13,32,36,39].

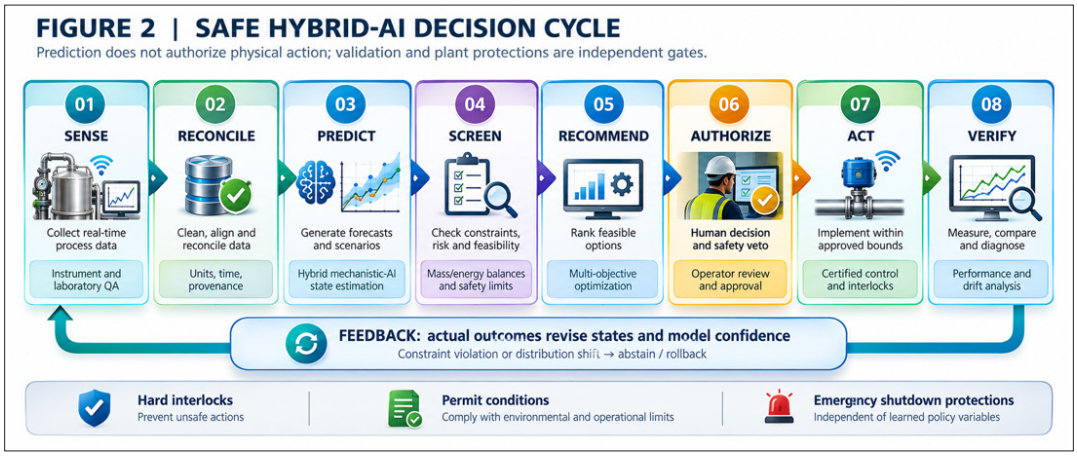


Figure 2: Sense-to-verify Decision Cycle Separating Advisory AI from Independent Safety Authorization (conceptual)

The control boundary is important for both safety and scientific evaluation. It is not reasonable for an offline model to directly overrule pressure interlocks, stop devices, gas alarms, or environmental discharge protection. Candidate actions should instead be selected from an operationally approved envelope and introduced at a frequency consistent with process dynamics and operator workload. Shadow-mode deployment is a valuable preliminary test because it records the recommendations a model would have made, permitting comparisons with actual outcomes without altering production. Yet a shadow-model's counterfactual performance remains estimated, so causal benefits ultimately require prospective, carefully monitored interventions [22,37,39,40].

A recent full-scale municipal anaerobic-digestion study used a calibrated digital twin for offline supervisory policy learning and counterfactual evaluation against historical plant operation. Its estimated improvement was modest and highly uncertain, and the learned policy was not tested through live plant intervention. This is useful evidence of control-oriented feasibility in an adjacent wastewater setting, not an observed causal improvement at a POME mill; it reinforces the need for prospective trials and explicit model–plant mismatch analysis [55].

Evaluation Beyond R^2: A Proposed Multidimensional Scorecard

Methodological comparison needs at least five dimensions. First, point prediction should be assessed using appropriate out-of-sample errors for methane, CO_2, H_2S, COD, and other measured states. Second, physical consistency should be quantified through mass and energy residuals, negative-concentration incidents, and violations of plausible rate bounds. Third, robustness should test temporal holdouts, unusually high loading, delayed data, sensor replacement, and truly new plants. Fourth, decision relevance should record avoided operating losses and the cost of missed alarms or false interventions. Fifth, practicability should assess inference latency, maintenance burden, data access, operator acceptance, and total life-cycle cost. No single statistical metric can substitute for these distinct tests [13,15,18,30,31,52].

The scorecard must account for how experimental evidence was produced. Some POME studies combine industrial historical records with statistical or software-based optimization, whereas the optimized settings may not have been enacted under randomized or independently controlled field conditions. Likewise, improved laboratory or batch digestion performance from adding biochar does not mean the same gain will occur continuously in a covered lagoon with changing feedstock and full-scale mixing constraints. Any comparative article should identify the evaluation scale, number of independent plants, chronological validity of test data, and whether a reported improvement refers to measured operation, a simulation, or a modeled counterfactual. This evidence grading protects the value of promising studies without overstating their conclusions [5,7,17,23,49,52].

Sim-to-Plant Transfer, Data Lineage, and Uncertainty-Aware Abstention

Cross-mill generalization requires greater discipline than transferring network weights. Mills differ in feed dilution, crop seasonality, residence-time distributions, pretreatment, lagoon geometry, microbial communities, gas collection, measurement frequency, and operational practice. A model validated on two plants therefore has stronger external support than one fitted to a single small dataset, but it is not automatically representative of all digesters or treatment technologies. A credible transfer protocol begins with a minimal common vocabulary for feed composition, mass loading, retention conditions, temperature, pH, and output units. It then tests zero-shot performance, local recalibration needs, and the stability of inferred kinetics before permitting operational use [8,15,19,22,35].

Missing laboratory data and concept drift should be treated as normal risks. A calibrated soft sensor for COD can become unreliable when feed constituents change or when an instrument is relocated; an H_2S predictor can degrade after scrubber media replacement or a change in gas composition. Uncertainty envelopes should include observation noise, model disagreement, parameter uncertainty, and lack of support in the historical data. Abstention is a legitimate system output: where the credible range of outcomes crosses a safety or environmental limit, the advisory layer should

decline to optimize and direct staff toward measurement or established conservative operation. Recorded abstentions, near misses, and overrides should become part of the model's governance dataset [18,30,31,36,53,54].

Recent chemical-engineering digital-twin studies have tested online updating and transfer learning, showing methodological routes for coping with changing process parameters and incomplete observability. Other wastewater work examines physics-informed effluent predictions, while digestion optimization research demonstrates trade-offs between prediction, treatment performance, and cost. These contributions motivate new POME-specific benchmarks; because their feedstocks and plant configurations differ, their reported advantages should not be transplanted as universal palm-mill performance estimates [14,52,55].

Implementation Roadmap, Organizational Ownership, and Evidence Gates

Figure 3 proposes an evidence-gated roadmap that begins with operational foundations rather than autonomous controllers. Phase I maps wastewater and residue streams, cleans historical datasets, audits sensors, and establishes laboratory validation. Phase II calibrates mechanistic and hybrid models, tests balances, and conducts chronological and cross-mill validation. Phase III runs shadow recommendations under independent safety review and compares them with conservative baselines. Phase IV permits narrowly bounded, reversible, human-authorized trials with prospective outcome measurement. At each gate, operators and environmental specialists review the evidence before authority is expanded [19,22,37,40,56].

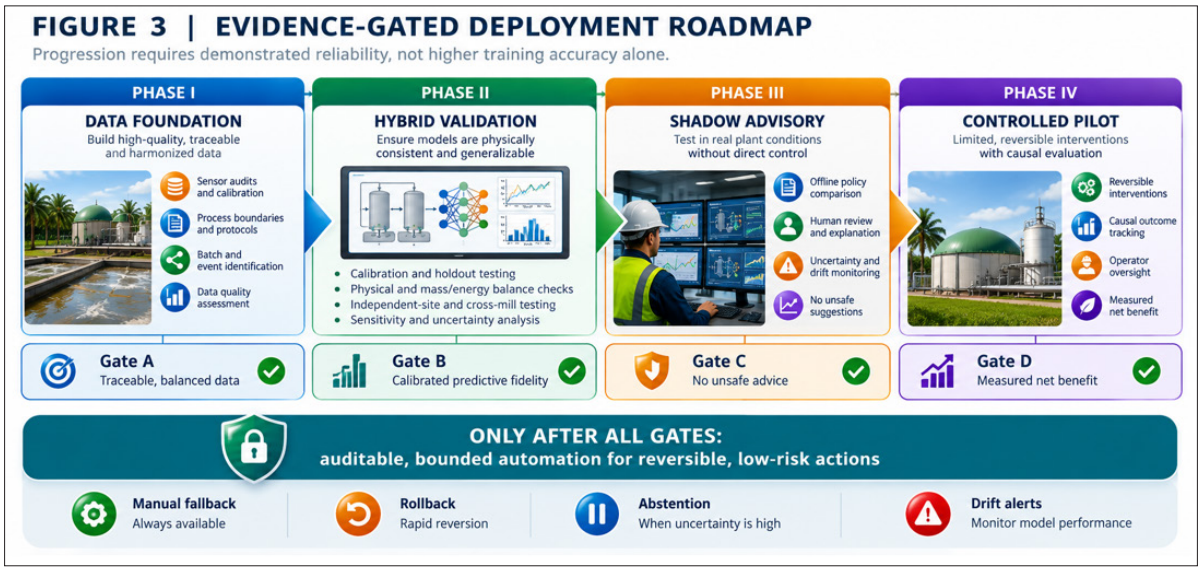


Figure 3: Evidence-gated implementation and validation roadmap for circular POME intelligence (proposed)

The roadmap reserves direct actuator responsibility for the plant's authorized control infrastructure and specifies rollback, manual fallback, model versioning, and incident escalation at every stage. The earliest use cases are likely to be soft sensing, abnormal-operation alerts, feed blending advice, and gas-routing recommendations, where existing sensor data and decision records support observable feedback. More ambitious multi-output optimization involving water reclamation, methane purification, and residue pyrolysis should follow only when subsystem balances and economic boundaries can be audited. The proposed ordering reduces the temptation to deploy complicated neural models before the mill can establish whether the underlying data and process assumptions are correct [20,24,37,57].

Synthesis Across the Research Questions and Unresolved Contradictions

The answer to RQ1 is that defensible physical guidance begins with explicit flows, inventories, feasible state bounds, and calibrated reduced digestion kinetics; not every microbial mechanism needs to be hard-coded. RQ2 receives a conditional answer: hybrid models may improve data efficiency and generalization where physics is informative and correctly specified, but they require fair competition with strong mechanistic and standard machine-learning baselines. RQ3 is answered by an architecture of quality-controlled sensing, bounded state estimation, constrained recommendations, external safety checks, authorization, and verified outcomes. These design choices follow from the accumulated literature without demonstrating that the complete architecture already operates at commercial palm oil mills [10,13,15,32,40,56].

For RQ4, the synthesis establishes that methane production, gas cleaning, biochar, water reuse, recovered oil, residual treatment, and energy allocation constitute an interacting resource system with potentially conflicting objectives. RQ5 indicates that credible implementation requires prospective testing, physical-balance audits, cross-site validation, full operating-cost and environmental accounting, cybersecurity, and operator decision rights. The most consequential unresolved contradiction is between high reported historical prediction accuracy and sparse evidence of long-duration, independently verified closed-loop improvement. Future research should therefore prioritize transparent multi-plant benchmarks, negative results, and causal impact evaluation over another incremental gain in training-set metrics. A circular biorefinery becomes meaningfully intelligent when its process benefits can be measured and its unsafe recommendations can be reliably prevented [1,4,19,21,58,62].

Conclusion and Recommendations

The review finds that physics-informed and hybrid artificial intelligence provides a credible approach to modeling complex POME digestion, especially where mechanistic balances constrain a data-driven component and difficult-to-measure states can be

estimated from reliable observations. The principal opportunity is not automatic maximization of methane production but reliable coordination of biological stability, recoverable energy, gas purification, water treatment, and solid-residue utilization. Such coordination requires a distinction between exact physical accounting, uncertain reaction mechanisms, predictive approximations, and independently enforceable operating limits. The Palm Circular Biorefinery Digital Intelligence framework developed here is a falsifiable conceptual synthesis, not a field-validated autonomous installation.

For researchers, the priority should be transparent, shared benchmark tasks and prospective evaluations that report error under shifted operating conditions, conservation residuals, calibration quality, failure frequency, operator burden, and actual economic and environmental performance. Comparisons must include competitive first-principles and conventional machine-learning baselines and should distinguish retrospective optimization from applied control. For mill companies, the appropriate sequence begins with provenance, data quality, laboratory measurement discipline, and model maintenance before escalating toward advisory optimization. Human oversight and independent safety interlocks should remain central whenever recommendations affect equipment, personnel, emissions, or wastewater discharge.

For sector organizations and policymakers, common data vocabularies, transparent reporting boundaries, and guidance for cross-mill validation would reduce the costs of trustworthy adoption. Environmental incentives should reflect verified methane capture, water quality, resource substitution, and lifecycle impacts rather than modeled gross recovery alone. The strategic conclusion is that a palm biorefinery becomes circular through accountable physical and economic flows, and becomes intelligent when its models improve those flows without obscuring uncertainty or transferring unacceptable risk to workers, communities, or the environment.

References

1. Manikandan SK (2026) Treating palm oil mill effluent (POME): Opportunities, challenges, and sustainable practices. Environmental Technology & Innovation 42: 104857.
2. Imam SS, Sani S, Mujahid M, Adnan R (2025) Valuable resources recovery from palm oil mill effluent (POME): A short review on sustainable wealth reclamation. Waste Management Bulletin 3: 1-16.
3. Tang YM, Tan KT, Wong LP (2024) Palm oil mill effluent (POME) as a source of biofuels and value-added products via oil recovery: a review. Journal of Oil Palm Research 36: 534-546.
4. Nasrin AB (2024) Renewable Energy and Greenhouse Gases Emission Reduction Potential of Biogas from Palm Oil Mill Effluent. Journal of Oil Palm Research 36: 456-468.
5. Tan VWG, Chan YJ, Arumugasamy SK, Lim JW (2023) Optimizing biogas production from palm oil mill effluent utilizing integrated machine learning and response surface methodology framework. Journal of Cleaner Production 414: 137575.
6. Chen JW, Chan YJ, Arumugasamy SK, Yazdi SK (2023) Process modelling and optimisation of methane yield from palm oil mill effluent using response surface methodology and artificial neural network. Journal of Water Process Engineering 52: 103493.
7. Yang PB, Chan YJ, Yazdi SK, Lim JW (2024) Optimisation and economic analysis of industrial-scale anaerobic co-digestion (ACoD) of palm oil mill effluent (POME) and decanter cake (DC) using machine learning models: A comparative study of Gradient Boosting Machines (GBM), K-nearest neighbours (KNN), and random forest (RF). Journal of Water Process Engineering 58: 104752.
8. Hoon YTJ, Chan YJ, Wan YK, Goh YK, Yazdi SK (2024) Industrial-scale anaerobic Co-digestion (ACoD) of palm oil mill effluent (POME) and decanter cake (DC) for maximizing methane yield: An integrated machine learning and simulation-based economic analysis approach. Energy 289: 129939.
9. Gupta R, Zhang L, Hou J, Zhang Z, Liu H (2023) Review of explainable machine learning for anaerobic digestion. Bioresource Technology 369: 128468.
10. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, et al. (2021) Physics-informed machine learning. Nature Reviews Physics 3: 422-440.
11. Mahanty B (2023) Hybrid modeling in bioprocess dynamics: Structural variabilities, implementation strategies, and practical challenges. Biotechnology and Bioengineering 120: 2072-2091.
12. Velioglu M, Zhai S, Rupprecht S, Mitsos A, Jupke A, et al. (2025) Physics-informed neural networks for dynamic process operations with limited physical knowledge and data. Computers & Chemical Engineering 192: 108899.
13. Koksal ES, Asrav T, Esenboga EE, Cosgun A, Kusoglu G, et al. (2024) Physics-informed and data-driven modeling of an industrial wastewater treatment plant with actual validation. Computers & Chemical Engineering 108801.
14. Li Z, Yu S, Yang B, Tang L, Wang N (2026) Reliable and interpretable effluent quality prediction in wastewater treatment plants through the integration of physics-informed neural network and deep learning. Journal of Environmental Chemical Engineering 14: 122957.
15. Shaw KM, Poh PE, Ho YK, Chen ZY, Chew IML (2024) Modeling the anaerobic digestion of palm oil mill effluent via physics-informed deep learning. Chemical Engineering Journal 485: 149826.
16. Gan EYT, Chan YJ, Wan YK, Tiong TJ, Chong WC, et al. (2024) Examining the synergistic effects through machine learning prediction and optimisation in the anaerobic Co-digestion (ACoD) of palm oil mill effluent (POME) and decanter cake (DC) with economic analysis. Journal of Cleaner Production 437: 140666.
17. Chang PY, Chan YJ, Arumugasamy SK, Wan YK, Lim JW (2025) Optimisation of anaerobic digestion of palm oil mill effluent with biochar addition: Synergistic application of Artificial neural network and response Surface Methodology. Fuel 398: 135514.
18. Clifford JL, Chan YJ, Mohd Yusof MA, Tiong TJ, Lim SS, et al. (2025) Predictive Modelling of H2S Removal from Biogas Generated from Palm Oil Mill Effluent (POME) Using a Biological Scrubber in an Industrial Biogas Plant: Integration of Artificial Neural Network (ANN) and Process Simulation. Food Technology and Biotechnology 63: 124-133.
19. Wang A-J, Li H, He Z, Tao Y, Wang H, et al. (2024) Digital Twins for Wastewater Treatment: A Technical Review. Engineering 36: 21-35.
20. Ho QN, Lau WJ, Jaafar J, Othman MHD, Yoshida N (2025) Membrane Technology for Valuable Resource Recovery from Palm Oil Mill Effluent (POME): A Review. Membranes 15: 138.
21. Goh KC (2025) Innovative circular bioeconomy and decarbonization approaches in palm oil waste management:

A review. Process Safety and Environmental Protection 195: 106746.

22. Murali R, Dekhici B, Du B, Guo B, Chen T, et al. (2026) Towards the integration of mechanistic, microbial, and machine learning approaches in anaerobic digestion modelling and control: a review of current progress and implementation strategies. Bioresource Technology: 135967.
23. Ge Y, Tao J, Wang Z, Mu L, Guo W, et al. (2024) A hybrid approach of anaerobic digestion model no. 1 and machine learning to model and optimize continuous anaerobic digestion processes. Biomass and Bioenergy 184: 107176.
24. Sulin SN, Mokhtar MN, Baharuddin AS, Mohammed MAP (2025) Simulation and techno-economic evaluation of integrated palm oil mill processes for advancing a circular economy. Cleaner Environmental Systems 19: 100323.
25. Soo PL, Bashir MJK, Wong L-P (2022) Recent advancements in the treatment of palm oil mill effluent (POME) using anaerobic biofilm reactors: Challenges and future perspectives. Journal of Environmental Management 320: 115750.
26. Dominic D (2022) Recent Developments in Biological Processing Technology for Palm Oil Mill Effluent Treatment-A Review. Biology 11: 525.
27. Yan SJ, Chan YJ, Thangalazhy-Gopakumar S, Tiong TJ, Lim JW (2025) Optimizing anaerobic digestion of palm oil mill effluent (POME) with biochar: Synergistic impact of biochar addition and kinetic analysis. Journal of Water Process Engineering 70: 106919.
28. Schweidtmann AM (2021) Machine Learning in Chemical Engineering: A Perspective. Chemie Ingenieur Technik 93: 2029-2039.
29. Rogers AW, Song Z, Vega Ramon F, Jing K, Zhang D (2023) Investigating greyness of hybrid model for bioprocess predictive modelling. Biochemical Engineering Journal 189: 108761.
30. Li J, Long X, Deng X, Jiang W, Zhou K, et at. (2024) A principled distance-aware uncertainty quantification approach for enhancing the reliability of physics-informed neural network. Reliability Engineering & System Safety 245: 109963.
31. Zou Z, Meng X, Karniadakis GE (2025) Uncertainty quantification for noisy inputs-outputs in physics-informed neural networks and neural operators. Computer Methods in Applied Mechanics and Engineering 433: 117479.
32. Luo J, Abdullah F, Christofides PD (2023) Model predictive control of nonlinear processes using neural ordinary differential equation models. Computers & Chemical Engineering 178: 108367.
33. Hajireza K, Eklund D, Andersson R (2025) Enhancing generalization and training efficiency of neural ordinary differential equations for chemical reactor modeling. Chemical Engineering Journal: 165183.
34. Albino M, Gargalo CL, Nadal-Rey G, Albæk MO, Krühne U, et al. (2024) Hybrid Modeling for On-Line Fermentation Optimization and Scale-Up: A Review. Processes 12: 1635.
35. Riezzo L, Kay H, Feng Y, Jing K, Zhang D (2025) Accelerating bioprocess digital twin development by integrating hybrid modelling with transfer learning. Chemical Engineering Journal 511: 162018.
36. Haimi H, Awaitey A, Kiran A, Larsson T, Blomberg K et al. (2025) Integrating data-driven models and process expertise in soft-sensor design for a wastewater treatment digital twin application. Water Science and Technology 92: 1308-1327.
37. Ma Z, Zhu Y, Chen C, Li T, Li Y, et al. (2025) Towards the digitalization of water treatment facilities: A case study on machine learning-enabled digital twins. Journal of Water Process Engineering 77: 108316.
38. Polak J, Huang Z, Sokolov M, von Stosch M, Butté A, et al. (2024) An innovative hybrid modeling approach for simultaneous prediction of cell culture process dynamics and product quality. Biotechnology Journal 19: e2300473.
39. Faria RdR, Capron BDO, Secchi AR, de Souza MB (2022) Where Reinforcement Learning Meets Process Control: Review and Guidelines. Processes 10: 2311.
40. Gu S et al. (2024) A Review of Safe Reinforcement Learning: Methods, Theories, and Applications. IEEE Transactions on Pattern Analysis and Machine Intelligence 46: 11216-11235.
41. Nasir FB, Li J (2024) Understanding machine learning predictions of wastewater treatment plant sludge with explainable artificial intelligence. Water Environment Research 96: e11136.
42. Nugroho FG (2025) Utilizing Indonesian Empty Palm Fruit Bunches: Biochar Synthesis via Temperatures Dependent Pyrolysis. Nanomaterials 15: 50.
43. Aryanti PTP, Harsono B, Biantoro MFW, Romariyo R, Putri TA, et al. (2025) The role of membrane technology in palm oil mill effluent (POME) decontamination: Current trends and future prospects. Journal of Environmental Management 374: 124094.
44. Mahmod SS, Takriff MS, AL-Rajabi MM, Abdul PM, Gunny AAN, Silvamany H, Md Jahim J (2023) Water reclamation from palm oil mill effluent (POME): Recent technologies, by-product recovery, and challenges. Journal of Water Process Engineering 52: 103488.
45. Djarot IN, Pawignya H, Handayani T, Widyastuti N, Isharyadi F, et al. (2024) Enhancing sustainability: microalgae cultivation for biogas enrichment and phycoremediation of palm oil mill effluent - a comprehensive review. Environmental Pollutants and Bioavailability 36: 2347314.
46. Tzachor A, Richards CE, Jeen S (2022) Transforming agrifood production systems and supply chains with digital twins. npj Science of Food 6: 47.
47. Melesse TY, Franciosi C, Di Pasquale V, Riemma S (2023) Analyzing the Implementation of Digital Twins in the Agri-Food Supply Chain. Logistics 7: 33.
48. Purlis E (2024) Digital Twin Methodology in Food Processing: Basic Concepts and Applications. Current Nutrition Reports 13: 914-920.
49. Ling JYX, Chan YJ, Chen JW, Chong DJS, et al. (2024) Machine learning methods for the modelling and optimisation of biogas production from anaerobic digestion: a review. Environmental Science and Pollution Research 31: 19085-19104.
50. Herrera-Ruiz JF, Fontalvo J, Prado-Rubio OA (2024) Advances on hybrid modelling for bioprocesses engineering: Insights into research trends and future directions from a bibliometric approach. Results in Engineering 24: 103548.
51. Bayer B, Duerkop M, Pörtner R, Möller J (2023) Comparison of mechanistic and hybrid modeling approaches for characterization of a CHO cultivation process: Requirements, pitfalls and solution paths. Biotechnology Journal 18: e2200381.
52. Karami A, Karimi-Jashni A, Nikoo MR, Zare B (2025) Simulation of anaerobic digestion process under variable feeding sludge using a hybrid machine learning model. Chemosphere 387: 144672.
53. Kwon S, Kang Y, Nam SH, Kim YD (2025) Water quality monitoring using hybrid physical-soft sensors for river digital twins: a comprehensive review. Water Science and

Technology 92: 1286-1307.
54. Rebello CM, Nogueira IBR (2025) Digital twins in chemical engineering: An integrated framework for identification, implementation, online learning, and uncertainty assessment. Computers & Chemical Engineering 200: 109178.
55. Karami A, Karimi-Jashni A, Nikoo MR, Zare B (2026) Hybrid machine learning–NSGA-II framework for multi-objective optimization of anaerobic digestion in wastewater treatment. Scientific Reports, advance online publication, 19 September.
56. Yunus AI, Jalla S, Bozeman JF, Chen Y (2026) Reinforcement-learning-enabled supervisory control of a full-scale anaerobic digestion facility using a calibrated digital twin. Environmental Science & Technology 60: 28064-28078.
57. Kallenberg MGJ (2025) Interoperable agricultural digital twins with reinforcement learning intelligence. Smart Agricultural Technology 12: 101412.
58. Guidani B, Ronzoni M, Accorsi R (2024) Virtual agri-food supply chains: A holistic digital twin for sustainable food ecosystem design, control and transparency. Sustainable Production and Consumption 46: 161-179.
59. Peterson L, Gosea IV, Benner P, Sundmacher K (2025) Digital twins in process engineering: An overview on computational and numerical methods. Computers & Chemical Engineering 193: 108917.
60. Verdouw C, Tekinerdogan B, Beulens A, Wolfert S (2021) Digital twins in smart farming. Agricultural Systems 189: 103046.
61. Liu M, Fang S, Dong H, Xu C (2021) Review of digital twin about concepts, technologies, and industrial applications. Journal of Manufacturing Systems 58: 346-361.
62. Monteiro J, Barata J (2025) Digital twin-enabled regional food supply chain: A review and research agenda. Journal of Industrial Information Integration 45: 100851.

Copyright: ©2026 Loso Judijanto. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.