The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 3, Issue 12, 2025

MACHINE LEARNING–BASED OPTIMIZATION OF NANOCATALYST PROPERTIES FOR ENHANCED BIODIESEL PRODUCTION

Muhammad Hamza Rehman[*1], **Kiran Batool**[2], **Sanaara**[3],

[*1]*Governament College University Faisalabad,*
[2]*Department Of Biochemistry Bahauddin Zakariya University Multan*
[3]*BS Chem*

[*1]hamzarehman908@yahoo.com, [2]kiranbatool1257@gmail.com, [3]sanaara543@gmail.com

DOI: https://doi.org/10.5281/zenodo.18239470

Keywords
Biodiesel; Nanocatalyst; Heterogeneous catalysis; Transesterification; Machine learning; Optimization; Catalyst design

Article History
Received 03 November 2025
Accepted 17 December 2025
Published 31 December 2025

Corresponding Author: *
Muhammad Hamza Rehman

Abstract

Biodiesel production through transesterification and esterification reactions is strongly influenced by catalyst chemistry, surface structure, and stability, particularly when low-cost feedstocks such as waste cooking oil and non-edible oils are employed. Nanostructured heterogeneous catalysts have gained significant attention due to their high surface area, tunable acid–base functionality, and potential for catalyst recovery and reuse. However, optimizing nanocatalyst properties remains challenging because catalyst composition, morphology, surface chemistry, and synthesis conditions interact nonlinearly with process variables and feedstock characteristics. In recent years, machine learning (ML) has emerged as a powerful tool for modeling complex reaction systems and optimizing biodiesel production parameters. While many ML studies focus on predicting biodiesel yield using operating conditions alone, fewer efforts explicitly integrate nanocatalyst physicochemical properties into data-driven optimization frameworks. This review critically examines recent advances in ML-assisted biodiesel production with an emphasis on the optimization of nanocatalyst properties. The discussion covers major nanocatalyst classes, catalyst descriptors relevant to ML modeling, commonly applied ML algorithms, validation strategies, and optimization techniques including metaheuristics and Bayesian approaches. Key challenges such as data scarcity, catalyst deactivation, reproducibility, and model generalization are highlighted. Finally, future research directions are proposed toward catalyst property–centric modeling, active learning, and inverse catalyst design for robust and sustainable biodiesel production.

INTRODUCTION

This review is designed to critically examine the role of machine learning in the optimization of nanocatalyst properties for biodiesel production, rather than providing a general overview of biodiesel processes or artificial intelligence techniques. The scope is intentionally focused on heterogeneous and nanostructured catalysts used in esterification and transesterification reactions, with particular attention to how catalyst physicochemical properties are represented, modeled, and optimized using data-driven approaches.

The literature considered in this review spans foundational biodiesel studies and recent advances in nanocatalysis and machine learning. Early reviews and process studies were included to establish the chemical and catalytic background of biodiesel production and to identify recurring challenges related to catalyst activity, stability, and feedstock variability (Ma and Hanna, 1999; Meher et al., 2006; Lam et al., 2010). More recent studies were prioritized to capture advances in nanocatalyst design, ML algorithms, and hybrid optimization frameworks, particularly those published within the last 10–12 years.A structured literature search strategy was adopted to ensure coverage and relevance. Peer-reviewed journal articles were identified primarily through major scientific databases, including Scopus, Web of Science, and ScienceDirect. Search keywords combined terms related to biodiesel chemistry ("biodiesel," "transesterification," "esterification," "FAME"), catalyst design ("nanocatalyst," "heterogeneous catalyst," "CaO," "solid acid," "magnetic catalyst"), and data-driven methods ("machine learning," "artificial neural network," "random forest," "genetic algorithm," "Bayesian optimization"). Review articles were used to map research trends and identify influential experimental studies, while original research papers were examined in detail to extract catalyst descriptors, modeling approaches, and optimization outcomes (Helwani et al., 2009; Kouzu and Hidaka, 2012; Awogbemi et al., 2023).

Inclusion criteria for the review emphasized studies that satisfied at least one of the following conditions: (i) explicit use of nanostructured or nanoscale heterogeneous catalysts for biodiesel production; (ii) application of ML models to predict biodiesel yield, conversion, or related performance metrics; or (iii) integration of ML with optimization techniques such as genetic algorithms, particle swarm optimization, or Bayesian frameworks. Studies that focused solely on homogeneous catalysis, enzymatic routes without ML analysis, or purely theoretical AI discussions without experimental relevance were excluded to maintain chemical and practical relevance.

Special attention was given to how catalysts were characterized and reported. Studies that included detailed physicochemical characterization—such as surface area, pore structure, acid/base site density, crystallinity, and reusability—were prioritized because these features enable meaningful ML-based generalization across catalyst systems (Tran et al., 2017; Lee and Wilson, 2015). Where available, information on catalyst deactivation, leaching, and reuse cycles was also considered, as these aspects are critical for assessing industrial applicability but are often underrepresented in ML-driven studies (Ishola et al., 2024; Osman et al., 2024).

Finally, the reviewed literature was analyzed thematically rather than chronologically. Studies were grouped according to catalyst class, type of ML model, and optimization strategy. This approach enables comparison across different catalyst systems and highlights gaps where ML has not yet been fully leveraged to address persistent chemical and process challenges in biodiesel production.

2. Scope and Review Methodology

Biodiesel production is fundamentally governed by the chemistry of triglycerides and free fatty acids reacting with short-chain alcohols to form fatty acid alkyl esters. The dominant industrial pathway is transesterification, in which triglycerides react with methanol or ethanol to produce fatty acid methyl or ethyl esters and glycerol as a by-product. This reaction proceeds through a sequence of reversible steps involving diglyceride and monoglyceride intermediates, and its overall rate and equilibrium are strongly influenced by catalyst type, alcohol excess, temperature, and mass transfer limitations (Ma and Hanna, 1999; Leung et al., 2010).

For refined vegetable oils with low free fatty acid (FFA) content, base-catalyzed transesterification is typically sufficient to achieve high conversion under mild conditions. However, when feedstocks contain elevated FFAs, as is common for waste cooking oil and animal fats, esterification becomes a necessary complementary reaction. Esterification converts FFAs into esters and water and is typically catalyzed by acidic catalysts. In the presence of strong base catalysts, FFAs readily form soaps, which consume catalyst, reduce ester yield, and complicate phase separation (Lam et al., 2010; Phan and Phan, 2008). Consequently, the chemical nature of the feedstock dictates not only catalyst selection but also the overall process configuration.From a process perspective, biodiesel production is a multiphase system involving immiscible oil and alcohol phases, with reaction occurring at phase boundaries or within catalyst pores in heterogeneous systems. Mass transfer resistance is therefore a critical factor, particularly in solid-catalyzed reactions where reactants must diffuse to active sites. Early studies emphasized the importance of mixing and alcohol-to-oil ratio in overcoming mass transfer limitations, while later work demonstrated that catalyst pore structure and surface chemistry play equally important roles (Helwani et al., 2009; Encinar et al., 2010).

Temperature also exerts a strong influence on reaction kinetics and equilibrium. Higher temperatures accelerate transesterification rates but increase energy consumption and may promote catalyst deactivation, leaching, or alcohol loss. For heterogeneous catalysts, operating temperatures must balance kinetic enhancement with catalyst stability. Solid base catalysts such as CaO, for example, exhibit high activity at moderate temperatures but can undergo surface hydration or carbonation if exposed to moisture or CO_2, leading to loss of basicity (Kouzu and Hidaka, 2012). Similarly, solid acid catalysts used for esterification may suffer from water inhibition because water is produced as a reaction product and competes for active sites (Lotero et al., 2005).

Catalyst stability and reusability are therefore fundamental process considerations. A catalyst that achieves high conversion in a single batch but rapidly deactivates or leaches active species may not be viable for repeated use. Numerous studies have reported declines in biodiesel yield over successive cycles due to surface fouling, structural changes, or loss of active components, particularly for waste-derived or unsupported oxide catalysts (Boey et al., 2011; Puspitasari et al., 2024). These effects are rarely captured by simple yield-based optimization and require explicit consideration in catalyst design and evaluation.

The choice of alcohol further influences reaction chemistry and process performance. Methanol is most commonly used due to its low cost and high reactivity, but ethanol offers advantages in terms of renewability and reduced toxicity. However, ethanol's lower polarity and larger molecular size can reduce reaction rates and exacerbate phase separation challenges, particularly in heterogeneous systems (Encinar et al., 2010). These trade-offs highlight the need for catalyst designs that are tailored to specific alcohols and feedstocks.

In heterogeneous catalysis, catalyst surface properties such as acidity/basicity, hydrophobicity, and pore accessibility strongly affect both esterification and transesterification pathways. Solid acids promote esterification by protonating the carbonyl group of FFAs, while solid bases facilitate transesterification through alkoxide formation. In bifunctional catalysts, these mechanisms may operate simultaneously, enabling one-pot processing of high-FFA feedstocks. However, improper balance or spatial arrangement of acid and base sites can lead to mutual neutralization and reduced activity (Lee and Wilson, 2015). Understanding these mechanistic aspects is essential for rational catalyst design.

From an optimization standpoint, biodiesel production chemistry presents a multivariable, nonlinear problem. Reaction kinetics, equilibrium limitations, mass transfer, and catalyst deactivation are interdependent and often feedstock-specific. Traditional optimization approaches that vary one factor at a time or rely on low-order polynomial models can fail to capture these complexities, particularly when catalyst properties are included as variables (Chakraborty and Sahu, 2014). This complexity provides strong motivation for data-driven modeling approaches that can learn patterns from experimental data without requiring explicit mechanistic equations.

Machine learning has therefore emerged as a complementary tool to classical chemical engineering analysis in biodiesel research. By incorporating process variables alongside feedstock descriptors and catalyst properties, ML models can approximate the underlying response surface and support efficient optimization. However, for ML predictions to be chemically meaningful and transferable, they must be grounded in an understanding of biodiesel reaction chemistry and process fundamentals. Without this foundation, ML-based optimization risks identifying operating conditions that are impractical, unstable, or incompatible with real-world feedstocks.

3. Biodiesel Production Chemistry and Process Fundamentals

Biodiesel production is fundamentally governed by the chemistry of triglycerides and free fatty acids reacting with short-chain alcohols to form fatty acid alkyl esters. The dominant industrial pathway is transesterification, in which triglycerides react with methanol or ethanol to produce fatty acid methyl or ethyl esters and glycerol as a by-product. This reaction proceeds through a sequence of reversible steps involving diglyceride and monoglyceride intermediates, and its overall rate and equilibrium are strongly influenced by catalyst type, alcohol excess, temperature, and mass transfer limitations (Ma and Hanna, 1999; Leung et al., 2010).

For refined vegetable oils with low free fatty acid (FFA) content, base-catalyzed transesterification is typically sufficient to achieve high conversion under mild conditions. However, when feedstocks contain elevated FFAs, as is common for waste cooking oil and animal fats, esterification becomes a necessary complementary reaction. Esterification converts FFAs into esters and water and is typically catalyzed by acidic catalysts. In the presence of strong base catalysts, FFAs readily form soaps, which consume catalyst, reduce ester yield, and complicate phase separation (Lam et al., 2010; Phan and Phan, 2008). Consequently, the chemical nature of the feedstock dictates not only catalyst selection but also the overall process configuration.

From a process perspective, biodiesel production is a multiphase system involving immiscible oil and alcohol phases, with reaction occurring at phase boundaries or within catalyst pores in heterogeneous systems. Mass transfer resistance is therefore a critical factor, particularly in solid-catalyzed reactions where reactants must diffuse to active sites. Early studies emphasized the importance of mixing and alcohol-to-oil ratio in overcoming mass transfer limitations, while later work demonstrated that catalyst pore structure and surface chemistry play equally important roles (Helwani et al., 2009; Encinar et al., 2010).

Temperature also exerts a strong influence on reaction kinetics and equilibrium. Higher temperatures accelerate transesterification rates but increase energy consumption and may promote catalyst deactivation, leaching, or alcohol loss. For heterogeneous catalysts, operating temperatures must balance kinetic enhancement with catalyst stability. Solid base catalysts such as CaO, for example, exhibit high activity at moderate temperatures but can undergo surface hydration or carbonation if exposed to moisture or CO_2, leading to loss of basicity (Kouzu and Hidaka, 2012). Similarly, solid acid catalysts used for esterification may suffer from water inhibition because water is produced as a reaction product and competes for active sites (Lotero et al., 2005).

Catalyst stability and reusability are therefore fundamental process considerations. A catalyst that achieves high conversion in a single batch but rapidly deactivates or leaches active species may not be viable for repeated use. Numerous studies have reported declines in biodiesel yield over successive cycles due to surface fouling, structural changes, or loss of active components, particularly for waste-derived or unsupported oxide catalysts (Boey et al., 2011; Puspitasari et al., 2024). These effects are rarely captured by simple yield-based optimization and require explicit consideration in catalyst design and evaluation.

The choice of alcohol further influences reaction chemistry and process performance. Methanol is most commonly used due to its low cost and high reactivity, but ethanol offers advantages in terms of renewability and reduced toxicity. However, ethanol's lower polarity and larger molecular size can reduce reaction rates and exacerbate phase separation challenges, particularly in heterogeneous systems (Encinar et al., 2010). These trade-offs highlight the need for catalyst designs that are tailored to specific alcohols and feedstocks.

In heterogeneous catalysis, catalyst surface properties such as acidity/basicity, hydrophobicity, and pore accessibility strongly affect both esterification and transesterification pathways. Solid acids promote esterification by protonating the carbonyl group of FFAs, while solid bases facilitate transesterification through alkoxide formation. In bifunctional catalysts, these mechanisms may operate simultaneously, enabling one-pot processing of high-FFA feedstocks. However, improper balance or spatial arrangement of acid and base sites can lead to mutual neutralization and reduced activity (Lee and Wilson, 2015). Understanding these mechanistic aspects is essential for rational catalyst design.

From an optimization standpoint, biodiesel production chemistry presents a multivariable, nonlinear problem. Reaction kinetics, equilibrium limitations, mass transfer, and catalyst deactivation are interdependent and often feedstock-specific.

Traditional optimization approaches that vary one factor at a time or rely on low-order polynomial models can fail to capture these complexities, particularly when catalyst properties are included as variables (Chakraborty and Sahu, 2014). This complexity provides strong motivation for data-driven modeling approaches that can learn patterns from experimental data without requiring explicit mechanistic equations.

Machine learning has therefore emerged as a complementary tool to classical chemical engineering analysis in biodiesel research. By incorporating process variables alongside feedstock descriptors and catalyst properties, ML models can approximate the underlying response surface and support efficient optimization. However, for ML predictions to be chemically meaningful and transferable, they must be grounded in an understanding of biodiesel reaction chemistry and process fundamentals. Without this foundation, ML-based optimization risks identifying operating conditions that are impractical, unstable, or incompatible with real-world feedstocks..

4. Nanocatalysts for Biodiesel: Types, Mechanisms, and Stability

Nanocatalysts represent a key advancement in heterogeneous catalysis for biodiesel production because they combine the advantages of solid catalysts with enhanced surface accessibility and tunable physicochemical properties. The nanoscale or nanostructured nature of these catalysts can reduce diffusion limitations, increase the fraction of exposed active sites, and enable control over surface chemistry that is difficult to achieve with bulk materials. In biodiesel systems, where large triglyceride molecules, immiscible phases, and water-sensitive reactions are common, these attributes are particularly important.

4.1 Basic oxide nanocatalysts

Basic oxide nanocatalysts are among the most widely studied systems for biodiesel transesterification. Calcium oxide (CaO) and magnesium oxide (MgO) exhibit strong basicity, which promotes the formation of alkoxide species necessary for nucleophilic attack on triglyceride carbonyl groups. At the nanoscale, these oxides provide higher external surface area and shorter diffusion paths compared with their bulk counterparts, often leading to higher apparent reaction rates (Kouzu and Hidaka, 2012).Waste-derived CaO nanocatalysts prepared from eggshells, mollusk shells, or other calcium-rich wastes are of particular interest because they combine low material cost with environmental benefits. Such catalysts have been shown to achieve high biodiesel yields from waste cooking oil under optimized conditions (Boey et al., 2011; Puspitasari et al., 2024). However, their performance is strongly influenced by preparation conditions such as calcination temperature and atmosphere, which determine crystallite size, surface basicity, and susceptibility to carbonation. Exposure to moisture and carbon dioxide can rapidly convert active CaO into less active $Ca(OH)_2$ or $CaCO_3$, leading to deactivation during storage or reuse. These stability issues highlight the importance of coupling catalyst characterization with performance evaluation across multiple cycles.

4.2 Acid nanocatalysts.

Acidic nanocatalysts play a critical role in biodiesel production from high-FFA feedstocks, where esterification must occur prior to or concurrently with transesterification. Solid acid catalysts promote esterification by protonating the carbonyl group of free fatty acids, facilitating nucleophilic attack by alcohol. Sulfonated carbon-based nanocatalysts, including sulfonated biochar and activated carbon, have gained prominence due to their high density of Brønsted acid sites and the possibility of deriving them from renewable biomass sources (Lotero et al., 2005; Yang et al., 2024).At the nanoscale, carbon-based acid catalysts can be engineered to balance hydrophilicity and hydrophobicity. This balance is crucial because esterification produces water as a by-product, which can inhibit acid sites and shift equilibrium. Hydrophobic surface regions can help repel water from active sites, maintaining activity under realistic conditions. However, excessive sulfonation or poor anchoring of $–SO_3H$ groups may lead to sulfur leaching, reducing catalyst lifetime and contaminating products. Therefore, stability and leaching resistance are central concerns for acid nanocatalysts, particularly when reuse is required.

4.3 Bifunctional nanocatalysts

Bifunctional nanocatalysts incorporate both acidic and basic functionalities within a single material, enabling simultaneous esterification of FFAs and transesterification of triglycerides. This one-pot processing approach is attractive for simplifying biodiesel production from low-quality feedstocks and reducing the number of processing steps. Mechanistically, acidic sites catalyze esterification, while basic sites drive transesterification reactions.However, designing effective bifunctional catalysts is nontrivial. Acid and base sites can neutralize each other if they are not spatially separated, leading to reduced overall activity. To address this, researchers have explored strategies such as core–shell architectures, compartmentalized porous structures, or selective functionalization of supports (Lee and Wilson, 2015). At the nanoscale, precise control over site distribution becomes possible but also more sensitive to synthesis conditions. As a result, reproducibility and scalability remain challenges for bifunctional nanocatalysts.

4.4 Magnetic nanocatalysts

Magnetic nanocatalysts introduce an additional functional dimension by enabling catalyst recovery through an external magnetic field. Typically based on iron oxide (Fe_3O_4) cores coated with catalytically active shells, these systems combine catalytic performance with operational convenience. Magnetic separation reduces filtration time and catalyst loss, making repeated reuse more feasible.Recent studies have demonstrated the successful application of magnetic sulfonated nanocatalysts for biodiesel production from mixed or non-edible feedstocks, often combined with ML-based optimization of process parameters (Ohale et al., 2025). Nevertheless, the stability of the magnetic core and the integrity of the catalytic shell under reaction conditions are critical. Shell degradation, iron leaching, or loss of magnetization can compromise both catalytic activity and separability over time.

4.5 Stability, deactivation, and reuse considerations

Across all nanocatalyst classes, stability and reusability are decisive factors for practical application. Catalyst deactivation can arise from multiple mechanisms, including leaching of active species, fouling by glycerol or organic residues, sintering of nanoparticles, and chemical transformation of active phases. In biodiesel systems, water and FFAs exacerbate many of these issues, particularly for basic oxide catalysts.Despite their importance, stability metrics are often underreported in experimental and ML-based studies. Many optimization efforts focus on maximizing initial yield without explicitly accounting for performance decay over multiple cycles. From a chemistry and engineering perspective, this omission limits the industrial relevance of reported “optimal” conditions. Studies that include reuse cycles and leaching analysis consistently show that catalyst performance can change significantly over time, underscoring the need to integrate stability descriptors into optimization frameworks (Puspitasari et al., 2024; Ishola et al., 2024).

4.6 Implications for ML-based optimization

The diversity of nanocatalyst types and deactivation mechanisms highlights why catalyst properties must be treated explicitly in ML models. Simply encoding catalyst identity or loading cannot capture differences in basicity, acidity, porosity, or stability. Instead, ML-based optimization should incorporate measurable catalyst descriptors that reflect underlying chemistry and structure. Doing so enables models not only to predict yield but also to identify trade-offs between activity and stability across catalyst classes.In summary, nanocatalysts offer powerful opportunities to enhance biodiesel production, but their successful deployment depends on careful control of composition, structure, and surface chemistry. Understanding the mechanistic roles and stability limitations of different nanocatalyst classes provides the necessary foundation for meaningful ML-based optimization, which is addressed in subsequent sections.

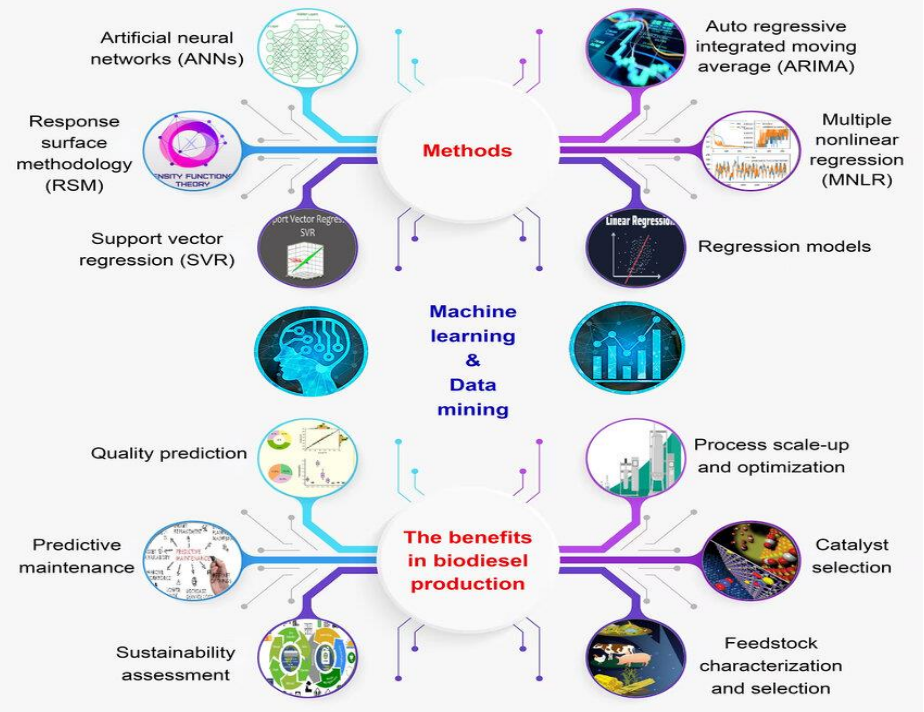

Figure 1. Machine learning–guided framework linking nanocatalyst properties, reaction conditions, and biodiesel performance.

5. Nanocatalyst Characterization and ML-Ready Descriptors

For machine learning to contribute meaningfully to nanocatalyst optimization, catalyst characteristics must be translated into quantitative, reproducible descriptors that reflect underlying chemistry and structure. In biodiesel catalysis, catalyst performance is governed not only by intrinsic activity but also by transport phenomena, stability, and resistance to deactivation. Therefore, ML-ready descriptors must capture multiple dimensions of catalyst behavior rather than relying solely on nominal composition or catalyst loading.

Table 1. Nanocatalyst descriptors relevant for ML-based biodiesel optimization

Descriptor Category	Examples	Importance in ML modeling
Textural	BET surface area, pore size	Controls reactant accessibility
Chemical	Acid/base site density	Determines reaction pathway
Structural	Crystallite size, phase	Affects stability
Morphological	Particle size, aggregation	Influences mass transfer
Stability	Leaching, reuse cycles	Industrial relevance

5.1 Textural and structural descriptors

Textural properties are among the most frequently reported catalyst characteristics and are critical for liquid-phase biodiesel reactions involving bulky triglyceride molecules. Specific surface area, pore volume, and pore size distribution—typically determined by nitrogen adsorption–desorption measurements—directly influence reactant accessibility and diffusion to active sites. Mesoporous structures are generally preferred for biodiesel catalysis because they balance high surface area with pore sizes large enough to accommodate triglycerides and diglyceride intermediates (Helwani et al., 2009; Tran et al., 2017).Structural descriptors obtained from X-ray diffraction (XRD), such as crystallite size and phase composition, provide insight into catalyst stability and active phase identity. For example, CaO catalysts with smaller crystallite sizes often exhibit higher initial activity but may also be more susceptible to hydration and carbonation, leading to faster deactivation (Kouzu and Hidaka, 2012). Including crystallite size as a continuous ML input enables models to capture trade-offs between activity and stability that are not apparent from catalyst identity alone.

5.2 Chemical and surface descriptors

Chemical descriptors are central to understanding catalytic function in biodiesel synthesis. For basic catalysts, surface basicity and base-site strength distributions determine the efficiency of alkoxide formation and subsequent transesterification steps. For acidic catalysts, Brønsted acid site density and strength govern esterification kinetics and tolerance to FFAs. These properties are commonly assessed using temperature-programmed desorption (TPD), titration methods, or spectroscopic techniques (Lotero et al., 2005; Lee and Wilson, 2015).Surface functional groups, such as sulfonic acid ($-SO_3H$) moieties on carbon catalysts, can be quantified through elemental analysis or spectroscopic methods. Including such descriptors in ML models allows differentiation between catalysts with similar surface areas but different chemical functionalities. Surface hydrophobicity, often inferred from contact angle measurements or indirectly from surface chemistry, is another important but underreported descriptor. Hydrophobic surfaces can reduce water adsorption on acid sites, thereby sustaining esterification activity under realistic reaction conditions (Yang et al., 2024).

5.3 Morphological descriptors

Morphological features such as particle size, particle size distribution, and degree of aggregation influence mass transfer and slurry behavior in heterogeneous biodiesel systems. Nanoparticles with severe aggregation may exhibit reduced effective surface area and hinder reactant diffusion, negating the advantages of nanoscale dimensions. Transmission and scanning electron microscopy are commonly used to assess these properties, but the resulting information is often reported qualitatively rather than quantitatively.From an ML perspective, approximate quantitative descriptors—such as average particle size or aggregation indices—are preferable to categorical descriptions. Including morphological descriptors helps ML models distinguish between catalysts that are chemically similar but physically different in ways that affect performance (Tran et al., 2017).

5.4 Stability and reusability descriptors

Catalyst stability is a decisive factor for industrial relevance but is frequently neglected in ML datasets. Stability-related descriptors include metal or functional group leaching (measured by ICP or elemental analysis), activity retention over reuse cycles, and changes in surface chemistry or structure after reaction. Studies that report biodiesel yield only for fresh catalysts provide limited guidance for real-world applications, where catalysts must operate over multiple cycles (Boey et al., 2011; Puspitasari et al., 2024).Incorporating stability metrics into ML models enables multi-objective optimization that balances yield with durability. For example, ML models could be trained to predict both initial conversion and conversion loss per cycle, thereby identifying catalysts that offer optimal long-term performance rather than peak short-term activity.

5.5 Linking synthesis parameters to descriptors

An important extension of descriptor-based modeling is the inclusion of synthesis parameters as ML inputs. Calcination temperature, impregnation conditions, sulfonation protocols, and activation treatments strongly influence catalyst structure and surface chemistry. By recording synthesis parameters alongside resulting catalyst descriptors, ML models can learn relationships between preparation routes and catalytic performance. This linkage is a prerequisite for inverse catalyst design, where ML suggests synthesis conditions likely to yield catalysts with desired properties (Ishola et al., 2024; Osman et al., 2024).

5.6 Descriptor selection and data quality considerations

Not all descriptors contribute equally to predictive performance, and overly large feature sets can lead to overfitting, particularly for small datasets. Feature selection and dimensionality reduction techniques can help identify the most informative descriptors. However, descriptor selection should be guided by chemical understanding rather than purely statistical criteria. Including chemically irrelevant or poorly measured variables can degrade model reliability and interpretability.Overall, the effectiveness of ML-based optimization in biodiesel catalysis depends critically on the quality, consistency, and relevance of catalyst descriptors. Comprehensive characterization, standardized reporting, and careful selection of ML-ready features form the bridge between experimental nanocatalyst development and data-driven optimization.

6. Machine Learning Models Used in Biodiesel Catalysis Studies

Machine learning models have been increasingly adopted in biodiesel research to capture the complex, nonlinear relationships between reaction conditions, feedstock properties, catalyst characteristics, and process performance. Unlike classical mechanistic or polynomial models, ML approaches do not require explicit assumptions about functional form, making them well suited for systems where multiple variables interact in non-additive ways. In biodiesel catalysis, these interactions are especially pronounced when heterogeneous nanocatalysts and variable-quality feedstocks are involved

6.1 Artificial neural networks and related models

Artificial neural networks (ANNs) are the most widely used ML models in biodiesel studies. ANNs consist of interconnected layers of neurons that learn nonlinear mappings between inputs and outputs through training. Early work demonstrated that ANN models could predict biodiesel yield with higher accuracy than response surface methodology (RSM), particularly when multiple operating parameters were varied simultaneously (Betiku and Adepoju, 2013; Chakraborty and Sahu, 2014). Subsequent studies applied ANN models to biodiesel production from various feedstocks, including waste cooking oil and non-edible oils, often achieving high predictive accuracy across broad experimental domains (Sivamani et al., 2019).Despite their strong predictive capability, ANN models present several limitations. Model performance depends heavily on network architecture, training algorithm, and data quality. Overfitting is a common risk, especially when datasets are small, which is frequently the case for nanocatalyst studies. Moreover, ANN models are often criticized for their limited interpretability, as the learned relationships are embedded within network weights that lack direct physical meaning. These limitations motivate the use of complementary modeling approaches and careful validation strategies.Adaptive neuro-fuzzy inference systems (ANFIS) combine neural networks with fuzzy logic, offering improved interpretability through rule-based structures. ANFIS models have been applied to biodiesel optimization and shown to perform comparably to ANN while providing clearer insight into variable interactions (Ighose et al., 2017). However, their scalability to large descriptor sets and high-dimensional catalyst data remains limited.

6.2 Support vector regression and kernel-based methods

Support vector regression (SVR) is another ML approach used in biodiesel modeling. SVR constructs a regression function by minimizing error within a defined margin while maximizing model generalization. Kernel functions enable SVR to capture nonlinear relationships without explicitly mapping data into high-dimensional spaces. SVR models have been successfully applied to biodiesel yield prediction and process optimization, particularly for moderate-sized datasets with well-defined input ranges (Betiku and Adepoju, 2013).However, SVR performance is sensitive to kernel choice and hyperparameter tuning. In complex nanocatalyst datasets that include diverse descriptor types, SVR may struggle to balance flexibility and generalization unless carefully optimized.

6.3 Tree-based ensemble models

Tree-based ensemble methods, such as random forest (RF) and gradient boosting algorithms, have gained increasing attention in recent biodiesel research. These models construct ensembles of decision trees to improve predictive accuracy and robustness. RF models are particularly attractive because they handle nonlinear interactions and mixed data types well while providing measures of variable importance that enhance interpretability.Recent studies applying RF and gradient boosting to biodiesel production from waste cooking oil have demonstrated improved predictive performance compared with ANN in some cases, especially when datasets include catalyst and feedstock descriptors alongside process variables (Ahmad et al., 2023). Gradient boosting methods further enhance accuracy by sequentially correcting prediction errors, although they require careful tuning to avoid overfitting.

6.4 Validation strategies and data leakage risks

Model validation is a critical yet often underemphasized aspect of ML studies in biodiesel catalysis. Many datasets include multiple experiments performed using the same catalyst under different conditions. Random train–test splits can therefore lead to data leakage, where information about a catalyst appears in both training and test sets, artificially inflating predictive performance.For nanocatalyst optimization, validation strategies should group data by catalyst batch or synthesis route, ensuring that test data represent genuinely unseen catalysts. This approach provides a more realistic assessment of a model's ability to generalize beyond known systems. Cross-validation, combined with multiple error metrics, further improves reliability and transparency.

6.5 ML versus RSM: complementary roles

Although ML models often outperform RSM in predictive accuracy, RSM retains value due to its simplicity and interpretability. In practice, ML and RSM can be complementary rather than competing approaches. RSM is useful for initial screening and experimental design, while ML models can capture higher-order interactions and support optimization in more complex domains (Chakraborty and Sahu, 2014; Ahmad et al., 2023).In summary, a diverse range of ML models has been applied to biodiesel catalysis, each with distinct strengths and limitations. Selecting an appropriate model requires careful consideration of dataset size, descriptor complexity, and the balance between predictive accuracy and interpretability. For nanocatalyst property optimization, models that can handle heterogeneous descriptors and support meaningful validation are particularly valuable.

7. ML-Based Optimization Strategies in Nanocatalyst-Assisted Biodiesel Production

Machine learning models are most impactful in biodiesel research when they are integrated with optimization strategies that actively guide experimental design and catalyst development. Rather than serving solely as predictive tools, ML models can function as surrogate models within optimization frameworks to identify combinations of catalyst properties and process variables that maximize biodiesel yield, stability, or overall process efficiency.

7.1 Metaheuristic optimization techniques

Metaheuristic algorithms such as genetic algorithms (GA) and particle swarm optimization (PSO) are among the most frequently used optimization techniques in ML-assisted biodiesel studies. These algorithms are inspired by natural or social processes and are well suited for exploring large, nonlinear search spaces without requiring gradient information. In biodiesel applications, ML models—often ANNs—are trained to predict biodiesel yield, and the trained models are then coupled with GA or PSO to search for optimal reaction conditions (Sivamani et al., 2019; Ighose et al., 2017).GA-based optimization has been widely applied to identify optimal temperature, alcohol-to-oil ratio, catalyst loading, and reaction time for biodiesel production from various feedstocks. These approaches can efficiently explore parameter interactions that are difficult to capture using traditional methods. However, GA and PSO are fundamentally deterministic once the surrogate model is fixed; they do not explicitly account for prediction uncertainty or experimental noise. As a result, optimized conditions may correspond to regions where the ML model is poorly informed by data.

7.2 Bayesian optimization and uncertainty-aware methods

Bayesian optimization has emerged as a more data-efficient alternative for optimizing biodiesel processes, particularly when experiments are costly or time-consuming. This approach treats the ML model as a probabilistic surrogate, often using Gaussian process regression, and explicitly incorporates uncertainty into the optimization process. By balancing exploration of uncertain regions with exploitation of promising conditions, Bayesian optimization can reduce the number of required experiments while improving robustness.In the context of nanocatalyst optimization, Bayesian methods are especially attractive because catalyst synthesis and characterization are resource-intensive. Bayesian optimization can guide experimental efforts toward catalyst compositions or synthesis conditions that are most likely to improve performance while also reducing uncertainty in poorly explored regions of the design space (Ishola et al., 2024; Osman et al., 2024). Although still relatively underutilized in biodiesel research, recent studies suggest strong potential for these methods to accelerate catalyst development.

7.3 Multi-objective optimization and trade-off analysis

Most biodiesel optimization studies focus on maximizing ester yield as the primary objective. However, from a practical perspective, catalyst stability, reusability, energy consumption, and environmental impact are equally important. Multi-objective optimization frameworks allow these competing objectives to be considered simultaneously, producing a set of Pareto-optimal solutions rather than a single "best" condition.In ML-assisted biodiesel research, multi-objective optimization can be implemented by training models to predict multiple outputs, such as yield and activity loss over reuse cycles, and then applying optimization algorithms to identify trade-offs. For example, a catalyst offering slightly lower initial yield but significantly improved stability may be preferable in industrial settings. Such trade-off analysis is rarely addressed in current ML-based studies but represents an important opportunity for advancing catalyst-centric optimization (Puspitasari et al., 2024).

7.4 Integration with experimental workflows

The ultimate goal of ML-based optimization is to create closed-loop experimental workflows in which models are continuously updated with new data. In such frameworks, ML models propose experimental conditions or catalyst designs, experiments are conducted, and the results are fed back into the model for retraining. This iterative process, often referred to as active learning, can significantly accelerate optimization while minimizing experimental effort.Although most biodiesel studies currently rely on static datasets, the integration of active learning with nanocatalyst experimentation represents a promising future direction. Implementing such workflows will require standardized data reporting, reliable characterization protocols, and careful consideration of uncertainty to ensure that ML-guided recommendations remain chemically meaningful.

Table 2. Representative ML-assisted biodiesel studies using nanocatalysts

Catalyst type	Feedstock	ML model	Optimization	Yield (%)
CaO nanoparticles	Waste cooking oil	ANN	GA	>95
Sulfonated biochar	Non-edible oil	ANN	RSM	˜92
Magnetic sulfonated catalyst	Mixed oils	XGBoost	Bayesian	˜96

Table 3. Comparison of ML techniques applied in biodiesel catalysis

ML method	Strengths	Limitations	Typical use
ANN	Strong nonlinear learning	Low interpretability	Yield prediction
Random Forest	Robust, interpretable	Large models	Screening
XGBoost	High accuracy	Complex tuning	Optimization
Gaussian Process	Uncertainty-aware	Computational cost	Bayesian optimization

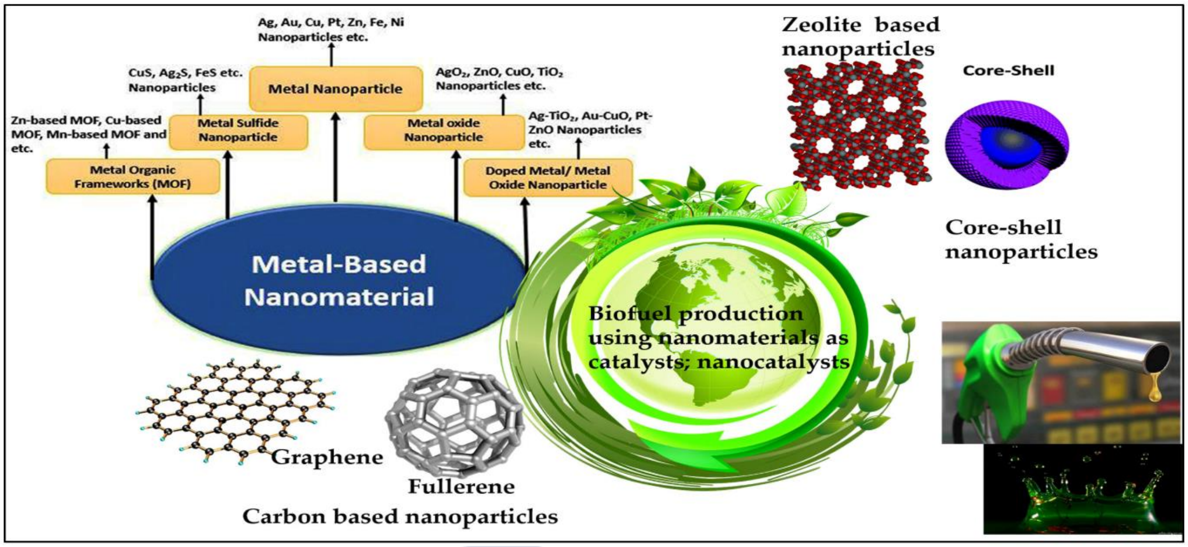

Figure 2. Structure–property–performance relationships in nanocatalyst-assisted biodiesel synthesis.

Analysis of Existing Machine Learning–Based Optimization Studies: Comparison and Discussion

Machine learning–based optimization has been increasingly applied to biodiesel production systems in recent years; however, the depth and focus of these applications vary considerably across studies. A critical analysis of existing literature reveals that most ML-assisted biodiesel studies emphasize process parameter optimization, while comparatively fewer investigations explicitly target **nanocatalyst physicochemical property optimization**. This distinction is important because process-centric models often lack transferability across catalyst systems, whereas property-centric models have greater potential to support rational catalyst design.Early applications of machine learning in biodiesel research primarily focused on predicting biodiesel yield as a function of operating variables such as reaction temperature, alcohol-to-oil ratio, catalyst loading, and reaction time. Artificial neural networks (ANN) coupled with response surface methodology (RSM) or genetic algorithms (GA) were commonly used and consistently outperformed traditional polynomial models in capturing nonlinear relationships (Betiku and Adepoju, 2013; Chakraborty and Sahu, 2014). While these studies demonstrated the superiority of ML over classical methods in terms of predictive accuracy, the catalyst itself was typically treated as a fixed entity or represented only by loading percentage, limiting broader applicability.More recent studies have attempted to incorporate catalyst descriptors into ML models, particularly when nanocatalysts are employed. Random forest, gradient boosting, and hybrid ANN-based models have been used to assess the relative importance of variables such as surface area, pore size, acidity, and particle size alongside process parameters (Ahmad et al., 2023). These studies highlight that catalyst-related features often exert a comparable or greater influence on biodiesel yield than operating conditions, underscoring the need to explicitly include nanocatalyst properties in ML-based optimization frameworks.A comparison of ML techniques reveals distinct strengths and limitations. ANN models are highly effective for nonlinear prediction but require careful architecture selection and large datasets to avoid overfitting. Support vector regression performs well for moderate datasets but is sensitive to kernel selection and parameter tuning. Tree-based ensemble methods, such as random forest and XGBoost, offer improved robustness and interpretability, particularly through feature-importance analysis, making them attractive for catalyst-property studies. However, these models may become computationally intensive as descriptor dimensionality increases.Another key difference among existing studies lies in their optimization objectives. Most investigations focus on maximizing biodiesel yield as a single objective, often neglecting catalyst stability, reusability, and leaching behavior. This limitation is particularly problematic for nanocatalysts, where surface restructuring, sintering, or functional group loss can significantly affect long-term performance (Boey et al., 2011; Puspitasari et al., 2024). Only a limited number of studies consider multi-objective optimization or include stability-related metrics in ML models, despite their importance for industrial relevanceValidation practices also vary widely across the literature. In many cases, random data splitting is used without accounting for shared catalyst identity or synthesis conditions, leading to potential data leakage and overly optimistic performance metrics. Studies employing grouped or catalyst-aware validation strategies provide more realistic assessments of model generalization but remain relatively rare. This inconsistency complicates direct comparison of reported ML performance across studies.Overall, the comparative analysis indicates that while machine learning has demonstrated clear advantages for biodiesel process optimization, its application to **nanocatalyst property optimization remains underdeveloped**. The most robust studies are those that integrate detailed catalyst characterization, employ ensemble or uncertainty-aware ML models, and adopt multi-objective optimization frameworks. Future work should prioritize standardized descriptor reporting, stability-inclusive objectives, and rigorous validation protocols to ensure that ML-based optimization moves beyond yield prediction toward genuine catalyst design and process scalability.

Aspect of Analysis	Conventional ML-Based Biodiesel Studies	Limitations Identified in Literature	Contribution of This Review
Optimization focus	Primarily process parameters (temperature, alcohol-to-oil ratio, catalyst loading)	Catalyst treated as fixed or black-box entity	Shifts focus toward **explicit nanocatalyst property optimization**
Catalyst representation in ML	Catalyst often encoded as type or loading only	Limited transferability across catalyst systems	Emphasizes **physicochemical descriptors** (surface area, acidity, stability)
Machine learning models used	ANN, SVR commonly applied	Overfitting risk and low interpretability	Highlights **ensemble and uncertainty-aware models** for catalyst datasets
Optimization strategy	Single-objective yield maximization	Ignores stability, reuse, and leaching	Advocates **multi-objective optimization** (yield–stability–energy trade-offs)
Treatment of catalyst stability	Rarely included in ML objectives	Poor industrial relevance	Identifies stability descriptors as **critical ML inputs**
Validation	Random data splitting	Risk of data leakage and	Recommends **catalyst-aware and**

Aspect of Analysis	Conventional ML-Based Biodiesel Studies	Limitations Identified in Literature	Contribution of This Review
practices		optimistic accuracy	**grouped validation**
Feedstock variability	Often implicit or neglected	Limited model generalization	Encourages inclusion of **feedstock descriptors**
Reusability considerations	Post-optimization or omitted	No long-term performance prediction	Integrates **reuse behavior into optimization framework**
Overall research orientation	Predictive, case-specific	Limited design guidance	Moves toward **property-centric and design-oriented ML frameworks**

8. Challenges, Data Gaps, and Reproducibility Issues

Despite the growing body of literature on machine learning–assisted biodiesel production, several challenges limit the reliability, transferability, and industrial relevance of reported optimization results. One of the most significant issues is the limited availability of high-quality, standardized datasets. Many biodiesel studies rely on small experimental datasets generated under narrowly defined conditions, which restricts the generalization capability of ML models and increases the risk of overfitting (Awogbemi et al., 2023; Ishola et al., 2024).

Another major challenge lies in inconsistent reporting of catalyst characterization and stability. While biodiesel yield is almost universally reported, essential catalyst descriptors such as acid/base site density, crystallite size, pore structure, and leaching behavior are often omitted or reported qualitatively. This inconsistency makes it difficult to compare studies or integrate data from multiple sources into unified ML frameworks (Tran et al., 2017; Osman et al., 2024). In particular, catalyst reusability and deactivation mechanisms are frequently neglected in ML-based optimization objectives, even though they are critical for assessing long-term process feasibility.

Reproducibility also represents a persistent concern. Variations in feedstock composition, especially for waste cooking oil, can significantly influence reaction outcomes. However, many studies treat feedstock properties as fixed or implicit variables rather than measurable inputs. As a result, ML models trained on one dataset may perform poorly when applied to different feedstocks or experimental setups (Kulkarni and Dalai, 2006; Phan and Phan, 2008).

From a modeling perspective, inappropriate validation strategies further undermine confidence in reported ML performance. Random data splitting without accounting for shared catalyst identity can lead to data leakage and overly optimistic error estimates. Without rigorous validation protocols, it is difficult to assess whether an ML model truly captures underlying chemical trends or merely memorizes experimental noise (Ahmad et al., 2023).

Addressing these challenges will require improved experimental design, standardized reporting practices, and greater emphasis on uncertainty-aware modeling. Without such efforts, ML-assisted optimization risks remaining a predictive exercise rather than a reliable tool for catalyst development and process scale-up.

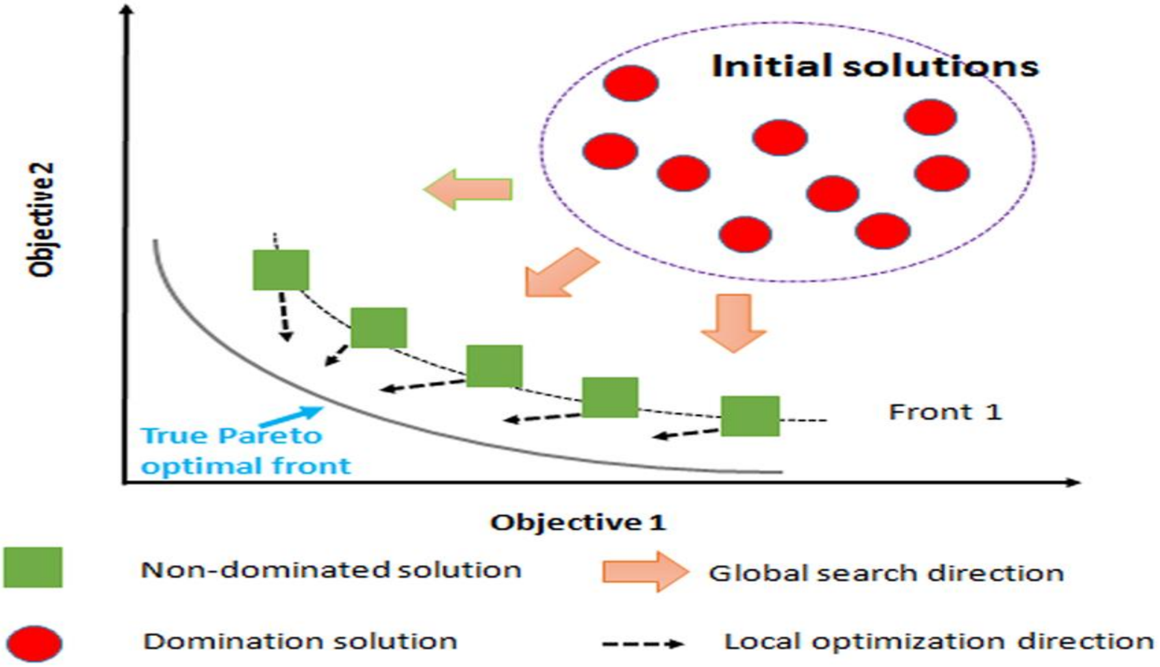

Figure 3. Conceptual Pareto optimization illustrating trade-offs between biodiesel yield, catalyst stability, and energy input.

9. Future Directions and Research Roadmap

Future research on machine learning–based optimization of nanocatalyst properties for biodiesel production should move beyond yield-centric modeling toward chemistry-informed, property-centric frameworks. One key direction is the systematic integration of catalyst synthesis parameters, physicochemical descriptors, and stability metrics into unified ML datasets. Such integration will enable models not only to predict performance but also to guide rational catalyst design rather than empirical parameter tuning (Ishola et al., 2024; Osman et al., 2024).Active learning and closed-loop optimization represent particularly promising avenues. In these approaches, ML models iteratively propose new catalyst compositions or operating conditions, experiments are performed, and the resulting data are used to update the model. This strategy can significantly reduce experimental effort while accelerating discovery, especially for nanocatalysts that require time-intensive synthesis and characterization.

Another important future direction is multi-objective optimization. Industrial biodiesel production demands trade-offs between yield, catalyst lifetime, energy consumption, and environmental impact. ML frameworks that simultaneously optimize multiple objectives can provide more realistic guidance for process development than single-objective yield maximization. Incorporating life-cycle and sustainability indicators into ML models may further enhance their relevance for policy and industrial decision-making.The development of open-access, community-shared datasets will also be critical. Standardized reporting of catalyst properties, feedstock composition, and experimental protocols would facilitate data reuse and improve model generalization across laboratories. Coupling such datasets with explainable ML techniques may help bridge the gap between predictive accuracy and chemical insight, fostering trust in data-driven recommendations.

Ultimately, the successful application of machine learning in biodiesel catalysis will depend on close integration between chemical understanding, high-quality experimental data, and robust modeling practices. By aligning ML tools with the realities of catalyst chemistry and process engineering, future studies can move toward genuinely predictive and design-oriented optimization of nanocatalysts for sustainable biodiesel production.

10. Conclusions

Machine learning–based optimization of nanocatalyst properties offers a powerful and flexible approach for improving biodiesel production under realistic feedstock and operating conditions. This review has shown that nanocatalysts, including basic oxides, solid acids, bifunctional systems, and magnetic nanocatalysts, provide important advantages in terms of surface accessibility, tunable chemistry, and potential reusability. However, their performance is governed by a complex interplay of physicochemical properties, synthesis conditions, and reaction parameters that cannot be efficiently optimized using conventional trial-and-error or low-order statistical methods alone.Machine learning models, when properly validated and supported by high-quality experimental data, can capture these nonlinear relationships and support efficient optimization. The review highlights that the true value of ML lies not only in predicting biodiesel yield but also in enabling property-centric and multi-objective optimization that accounts for catalyst stability, reusability, and process robustness. Persistent challenges related to data scarcity, inconsistent catalyst characterization, and limited reproducibility must be addressed for ML tools to become reliable design instruments. Overall, integrating nanocatalyst chemistry with data-driven modeling provides a clear pathway toward rational catalyst design and sustainable biodiesel production.

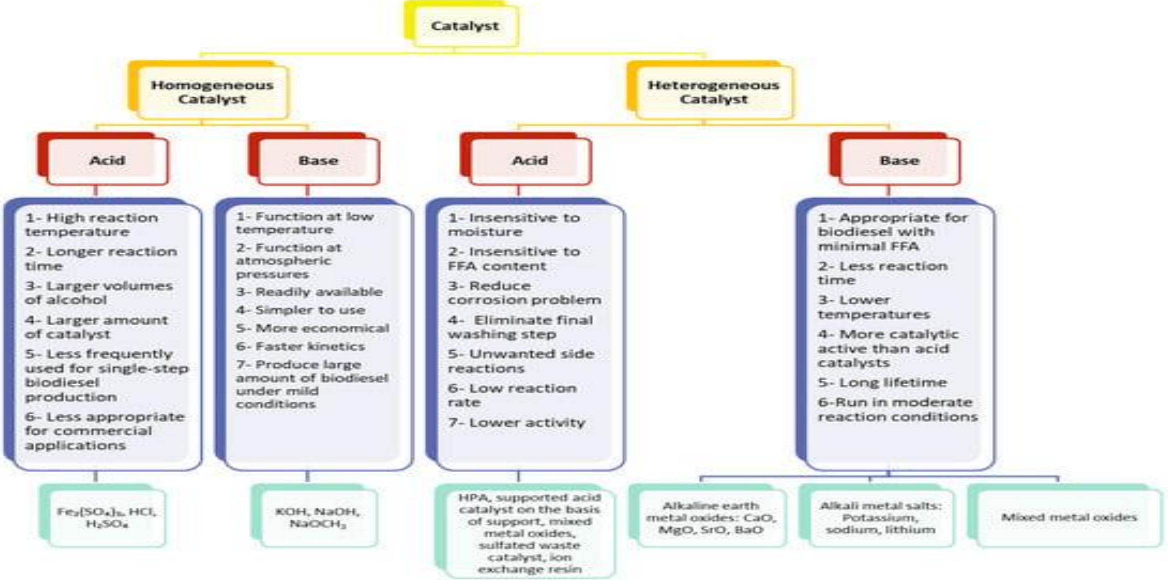

Figure 4. Major classes of nanocatalysts applied in biodiesel synthesis.

REFERENCES:

Ma, F.; Hanna, M.A. Biodiesel production: A review. *Bioresource Technology* **1999**, *70*, 1–15.

Meher, L.C.; Vidya Sagar, D.; Naik, S.N. Technical aspects of biodiesel production by transesterification—A review. *Renewable and Sustainable Energy Reviews* **2006**, *10*, 248–268.

Leung, D.Y.C.; Wu, X.; Leung, M.K.H. A review on biodiesel production using catalyzed transesterification. *Applied Energy* **2010**, *87*, 1083–1095.

Lam, M.K.; Lee, K.T.; Mohamed, A.R. Homogeneous, heterogeneous and enzymatic catalysis for transesterification of high free fatty acid oil. *Biotechnology Advances* **2010**, *28*, 500–518.

Di Serio, M.; Tesser, R.; Pengmei, L.; Santacesaria, E. From homogeneous to heterogeneous catalysts in biodiesel production. *Industrial & Engineering Chemistry Research* **2008**, *47*, 6379–6386.

Helwani, Z.; Othman, M.R.; Aziz, N.; Kim, J.; Fernando, W.J.N. Solid heterogeneous catalysts for transesterification of triglycerides with methanol: A review. *Applied Catalysis A: General* **2009**, *363*, 1–10.

Kouzu, M.; Hidaka, J. Transesterification of vegetable oil into biodiesel catalyzed by CaO: A review. *Fuel* **2012**, *93*, 1–12.

Boey, P.L.; Maniam, G.P.; Hamid, S.A. Utilization of waste cockle shell as catalyst for biodiesel production. *Bioresource Technology* **2011**, *102*, 364–368.

Betiku, E.; Adepoju, T.F. Modeling and optimization of biodiesel production process using artificial neural network. *Renewable Energy* **2013**, *53*, 177–185.

Chakraborty, R.; Sahu, H. Intensification of biodiesel production using infrared radiation: Process evaluation through response surface methodology and artificial neural network. *Applied Energy* **2014**, *114*, 827–836.

Sivamani, S.; Selvakumar, S.; Rajendran, K.; Muthusamy, S. Artificial neural network–genetic algorithm optimization of biodiesel production from *Simarouba glauca*. *Biofuels* **2019**, *10*, 393–401.

Ighose, B.O.; Damos, M.; Adeleke, I.A.; Junaid, H.A.; Okpalaeke, K.E.; Betiku, E. Optimization of biodiesel production from *Thevetia peruviana* seed oil by ANFIS coupled with GA and RSM. *Energy Conversion and Management* **2017**, *132*, 231–240.

Jayaraman, J.; Krishnan, R.; Subramaniam, D.; et al. Biodiesel production from waste cooking oil using ZnO nanoparticles combined with heteropoly acid catalyst. *Fuel* **2022**, *329*, 125411.

Naeem, A.; Khan, I.W.; Farooq, M.; et al. Biodiesel production from waste cooking oil using bentonite-supported heterogeneous catalyst. *Korean Journal of Chemical Engineering* **2022**, *39*, 1450–1459.

Yusuff, A.S.; Ishola, N.B.; Gbadamosi, A.O.; et al. Artificial intelligence modeling and optimization of biodiesel production from used cooking oil using CaO-incorporated zeolite catalyst. *Energy Conversion and Management: X* **2023**, *20*, 100452.

Awogbemi, O.; Dada, A.; Inambao, F. Application of machine learning technologies in biodiesel production process: A review. *Frontiers in Energy Research* **2023**, *11*, 1122638.

Ishola, N.B.; Epelle, E.I.; Betiku, E. Machine learning approaches to modeling and optimization of biodiesel production systems: State-of-the-art and future outlook. *Energy Conversion and Management: X* **2024**, *100669*.

Osman, A.I.; Fawzy, S.; Farghali, M.; et al. Optimizing biodiesel production from waste using computational chemistry and machine learning: A review. *Environmental Chemistry Letters* **2024**, *22*, 1–22.

Puspitasari, P.; Pramono, D.D.; Fiansyah, D.N.; et al. Biodiesel production from waste cooking oil using calcium oxide derived from scallop shell waste. *Clean Energy* **2024**, *8*, 113–126.

Ramalingam, K.; Arun, K.; Prabhakaran, S.; et al. Maximizing biodiesel production using machine learning and experiments with reusable CaO catalyst. *Scientific Reports* **2025**, *15*, 33265.

Ohale, P.E.; Amenaghawon, N.A.; Kimble Audu, T.O.; et al. Machine learning-informed one-pot biodiesel synthesis over magnetic sulfonated biobased catalyst. *RSC Advances* **2025**, *15*, 50856–50880.

Yang, Z.; Liu, X.; Wang, J.; et al. Efficient preparation of biodiesel using sulfonated *Camellia oleifera* shell biochar. *Molecules* **2024**, *29*, 2752.

Hossain, M.A.; Boyce, A.N. Biodiesel production from waste cooking oil using heterogeneous catalysts. *Renewable Energy* **2009**, *34*, 1503–1507.

Endalew, A.K.; Kiros, Y.; Zanzi, R. Heterogeneous catalysis for biodiesel production from vegetable oils. *Biomass and Bioenergy* **2011**, *35*, 3787–3809.

Banković-Ilić, I.B.; Stamenković, O.S.; Veljković, V.B. Biodiesel production from non-edible plant oils. *Renewable and Sustainable Energy Reviews* **2014**, *32*, 238–254.

Lotero, E.; Liu, Y.; Lopez, D.E.; et al. Synthesis of biodiesel via acid catalysis. *Industrial & Engineering Chemistry Research* **2005**, *44*, 5353–5363.

Lee, A.F.; Wilson, K. Structure–reactivity correlations in solid base catalyzed biodiesel synthesis. *Catalysis Today* **2015**, *242*, 3–18.

Tran, D.T.; Chang, J.S.; Lee, D.J. Recent insights into heterogeneous catalysts for biodiesel production. *Applied Energy* **2017**, *185*, 1562–1576.

Ahmad, A.; Yadav, A.K.; Singh, A. Application of machine learning and genetic algorithm to the prediction and optimization of biodiesel yield from waste cooking oil. *Korean Journal of Chemical Engineering* **2023**, *40*, 2941–2956.

Pillai, N.S.; Kannan, P.S.; Vettivel, S.C.; Suresh, S. Optimization of biodiesel production using green catalyst derived from biomass. *Renewable Energy* **2017**, *104*, 185–196.

Rashid, U.; Anwar, F. Production of biodiesel through optimized alkaline-catalyzed transesterification. *Fuel* **2008**, *87*, 265–273.

Granados, M.L.; Poves, M.D.Z.; Alonso, D.M.; et al. Biodiesel from sunflower oil by heterogeneous catalysis. *Applied Catalysis B: Environmental* **2007**, *73*, 317–326.

Encinar, J.M.; González, J.F.; Rodríguez-Reinares, A. Biodiesel from vegetable oils: Transesterification reaction. *Fuel* **2010**, *89*, 2211–2218.

Demirbas, A. Biodiesel fuels from vegetable oils. *Energy Sources* **2002**, *24*, 835–841.

Demirbas, A. Progress and recent trends in biodiesel fuels. *Energy Conversion and Management* **2009**, *50*, 14–34.

Phan, A.N.; Phan, T.M. Biodiesel production from waste cooking oils. *Fuel* **2008**, *87*, 3490–3496.

Kiss, A.A.; Dimian, A.C.; Rothenberg, G. Biodiesel production by reactive distillation. *Energy* **2010**, *35*, 598–606.

Van Gerpen, J. Biodiesel processing and production. *Fuel Processing Technology* **2005**, *86*, 1097–1107.

Zabeti, M.; Wan Daud, W.M.A.; Aroua, M.K. Activity of solid catalysts for biodiesel production. *Fuel Processing Technology* **2009**, *90*, 770–777.

Xie, W.; Li, H. Alumina-supported potassium iodide as a heterogeneous catalyst for biodiesel production. *Fuel* **2006**, *85*, 793–799.

Kouzu, M.; Kasuno, T.; Tajika, M.; et al. Calcium oxide as a solid base catalyst for biodiesel production. *Applied Catalysis A: General* **2008**, *334*, 357–365.

Deng, X.; Fang, Z.; Liu, Y. Biodiesel production from waste oils using solid catalysts. *Energy* **2011**, *36*, 777–784.

Liu, X.; He, H.; Wang, Y.; Zhu, S.; Piao, X. Transesterification of soybean oil to biodiesel using CaO as a solid base catalyst. *Fuel* **2008**, *87*, 216–221.

Marchetti, J.M.; Miguel, V.U.; Errazu, A.F. A review of biodiesel production processes. *Renewable and Sustainable Energy Reviews* **2007**, *11*, 1300–1311.

Vicente, G.; Martínez, M.; Aracil, J. Integrated biodiesel production: A comparison of different homogeneous catalysts. *Energy & Fuels* **2004**, *18*, 145–151.

Kulkarni, M.G.; Dalai, A.K. Waste cooking oil as an economical source for biodiesel. *Industrial & Engineering Chemistry Research* **2006**, *45*, 2901–2913.

Banković-Ilić, I.B.; et al. Technical aspects of biodiesel production from waste oils. *Renewable and Sustainable Energy Reviews* **2012**, *16*, 3621–3647.

Santoso, A.; et al. Ultrasonic-assisted biodiesel synthesis using heterogeneous catalysts. *AIMS Energy* **2022**, *10*, 824–838.

Veljković, V.B.; Stamenković, O.S.; Tasić, M.B. Biodiesel production: State of the art and perspectives. *Renewable and Sustainable Energy Reviews* **2011**, *15*, 145–156.

Bozbas, K. Biodiesel as an alternative motor fuel. *Energy Sources* **2008**, *30*, 963–975.

Ma, Y.; Wang, Q.; Sun, L.; et al. Advances in heterogeneous catalysis for biodiesel synthesis. *Catalysts* **2021**, *11*, 1154.

Li, X.; Zhang, H.; Chen, Y.; et al. Catalytic properties and structural optimization of solid catalysts for biodiesel synthesis. *Catalysts* **2025**, *15*, 239.

Atadashi, I.M.; Aroua, M.K.; Aziz, A.A. High quality biodiesel and its diesel engine application. *Renewable and Sustainable Energy Reviews* **2010**, *14*, 1999–2008.

Encinar, J.M.; Pardal, A.; Martínez, G. Transesterification of rapeseed oil with methanol. *Fuel* **2011**, *90*, 253–262.

Rashid, U.; Anwar, F.; Knothe, G. Evaluation of biodiesel obtained from different feedstocks. *Energy & Fuels* **2009**, *23*, 2927–2933