

Article

Multivariate Insights into Structure–Property Relationships in Biodiesel and Biolubricants from Vegetable Oils

Sergio Nogales-Delgado 

Department of Applied Physics, University of Extremadura, Avda. De Elvas s/n, 06006 Badajoz, Spain;
senogalesd@unex.es

Featured Application

In the context of biodiesel and biolubricants, PCA helps simplify complex datasets involving properties such as oxidation stability and viscosity. By reducing multiple correlated variables into a few key components, it becomes easier to identify how feedstock composition or processing conditions influence these critical performance parameters. It supports the selection and optimization of formulations, improving product quality while reducing experimental effort, and facilitating clearer interpretation of relationships between chemical composition and functional properties.

Abstract

The transition to bio-based fuels requires a thorough understanding of how molecular structure governs key physicochemical properties such as oxidation stability and viscosity. In this study, the combined influence of fatty acid composition and alcohol structure on biodiesel and biolubricant performance is investigated using a multivariate statistical approach. A dataset comprising 108 samples was analyzed, including 17 experimental samples produced in this work and 91 samples collected from peer-reviewed literature. Each sample was characterized by its fatty acid composition and at least one physicochemical property, namely oxidation stability (Rancimat induction time) and/or kinematic viscosity. Principal Component Analysis (PCA) was applied to identify dominant compositional patterns, followed by clustering (k-means) to define homogeneous compositional regions. Within these regions, a variance decomposition approach was used to quantify the relative contribution of alcohol type to property variability. The results show that fatty acid composition defines the primary structural framework governing oxidative stability, largely driven by the degree of unsaturation. In contrast, viscosity is more strongly influenced by the type of alcohol used, particularly in systems involving higher alcohols or polyols. The proposed approach provides a structured multivariate framework focused on analyzing oxidation stability and viscosity, enabling the systematic interpretation of the relative influence of fatty acid composition and alcohol structure on these key physicochemical properties. This study demonstrates that integrating PCA, clustering, and variance decomposition offers a robust strategy for analyzing complex bio-based systems, supporting the rational selection of feedstocks and processing conditions for the optimization of biodiesel and biolubricant formulations.



Academic Editor: Silvia Tabasso

Received: 15 May 2026

Revised: 9 June 2026

Accepted: 12 June 2026

Published: 15 June 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

Keywords: biofuels; lubrication; oleic acid; linoleic acid; linolenic acid; oxidative stability; viscosity; principal component analysis; PCA

1. Introduction

The shift toward a more sustainable energy and materials system has increased interest in renewable, biodegradable alternatives to petroleum-based products such as biodiesel and biolubricants [1,2]. Over the past decade (Figure 1), these bioproducts have gained strategic importance, influencing global energy dynamics, geopolitics, economic and environmental interdependencies [3–5].

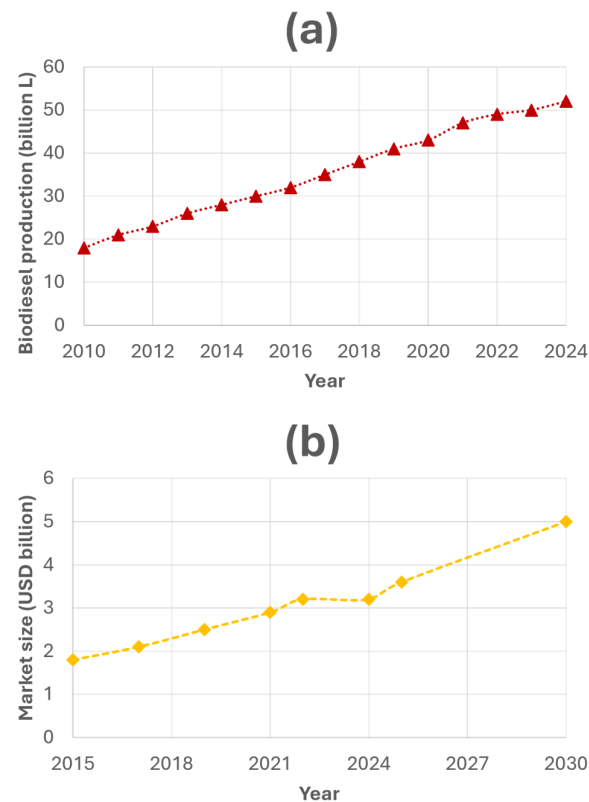


Figure 1. Global trends in (a) biodiesel production, in red triangles [6–10] and (b) biolubricant market size, in yellow squares [11–15]. Data compiled from international reports. Historical values correspond to reported data from international statistical and market sources. Projected values (e.g., for 2030) are derived from forecast analyses and are included solely to illustrate expected trends; they were not used in the statistical analyses of this study.

As observed in Figure 1a, global biodiesel production has increased steadily over the past decade, rising from approximately 18 billion liters in 2010 to around 50 billion liters in 2023, reflecting the growing role of bio-based fuels in the global energy transition [6–10].

Regarding Figure 1b, the global biolubricants market has experienced steady growth, reaching approximately USD 3 billion in 2024 and projected to exceed USD 5 billion by 2030, driven by increasing environmental regulations and demand for biodegradable alternatives to petroleum-based lubricants [11–15]. It should be noted that projected values included in Figure 1b are presented only to illustrate general trends and were not used in any of the statistical analyses conducted in this study.

Their growth is tied to the transition toward renewable energy, as countries seek to reduce fossil fuel dependence. Biodiesel has expanded rapidly due to policy mandates in the EU, the US, and Brazil, reshaping trade and strengthening producers like Brazil, Argentina, Indonesia, and Malaysia, though raising concerns about sustainability and trade practices [5,16].

Beyond fuels, biolubricants diversify supply chains by reducing reliance on petroleum and supporting agriculture-based value chains. In contexts such as Spain, biodiesel reflects

the balance between global competition and domestic resilience, linking energy and agricultural policies and influencing land use and markets. Thus, biodiesel and biolubricants have become strategic assets shaping energy policy, trade, and environmental goals [5].

Biodiesel and biolubricants are renewable, biodegradable products derived from vegetable oils, animal fats, or waste oils. Biodiesel consists mainly of fatty acid methyl esters (FAME) and offers favorable combustion and emission characteristics [16]. Biolubricants, typically based on esters or triglycerides, provide high lubricity, viscosity stability, and biodegradability, making them suitable for environmentally sensitive applications. Both benefit from the polarity of ester groups, which improves lubrication performance [17].

Their production relies on transesterification, where triglycerides react with alcohols to form esters and glycerol. This process, typically performed under mild conditions with alkaline catalysts, allows the tailoring of molecular structures to achieve desired viscosity and stability. Biodiesel is widely used in transportation and heating, while biolubricants are applied in industrial, hydraulic, and marine systems. Together, they demonstrate how renewable feedstocks can be converted into high-value products that support sustainable energy systems and circular economy strategies.

The performance of these bioproducts depends strongly on their physicochemical properties, which are governed by fatty acid composition. Key parameters such as viscosity or oxidation stability determine performance, safety, and compliance with standards (e.g., ASTM or EN). Viscosity affects flow and lubrication, while oxidation stability is critical due to the susceptibility of unsaturated fatty acids to degradation. High stability is essential to prevent the formation of degradation products that impair performance. These properties highlight the direct link between chemical structure and functional behavior [18].

The fatty acid composition of a given oil is shaped by biological, environmental, and processing factors, resulting in a wide diversity of lipid matrices with distinct structural features. Chain length distribution, degree of unsaturation, and the presence of functional groups or positional isomers all contribute to the reactivity and behavior of the resulting esters.

For instance, polyunsaturated fatty acids tend to accelerate oxidative degradation due to their susceptibility to radical-mediated reactions, whereas saturated fatty acids generally enhance oxidative resistance but may increase viscosity and reduce cold-flow performance. Similarly, monounsaturated fatty acids often provide a balance between stability and fluidity, making them desirable for certain applications.

Some of the most common fatty acids in vegetable oils for biodiesel and biolubricant production are palmitic acid (C16:0), a saturated fatty acid abundant in palm oil; stearic acid (C18:0), another saturated component found in smaller proportions in many seed oils; oleic acid (C18:1), a monounsaturated fatty acid dominant in oils such as canola, olive, and high-oleic sunflower; linoleic acid (C18:2), a polyunsaturated fatty acid prevalent in soybean and standard sunflower oils; and linolenic acid (C18:3), a highly unsaturated fatty acid present in linseed and some rapeseed varieties (see Figure 2) [19].

Saturated fatty acids like palmitic and stearic contribute to higher cetane numbers and better oxidative stability but worsen cold-flow properties, causing biodiesel to gel at lower temperatures. Monounsaturated oleic acid is considered the ideal compromise, offering good oxidative stability, favorable cold-flow behavior, and high combustion quality, which is why high-oleic feedstocks are increasingly preferred in industrial biodiesel production. Polyunsaturated fatty acids such as linoleic and linolenic improve cold-flow properties but significantly reduce oxidation stability, making the resulting biodiesel more prone to degradation during storage.

The relative proportions of these fatty acids determine key fuel parameters such as viscosity, iodine value, oxidation stability, and cold-filter plugging point, ultimately

shaping the commercial suitability of each feedstock. Understanding fatty-acid composition is therefore essential for selecting optimal raw materials, predicting biodiesel performance, and designing blends that meet international quality standards [20].

However, these relationships are not strictly linear, and the interplay among multiple fatty acid species can generate emergent effects that complicate straightforward prediction of product properties. This complexity underscores the need for analytical approaches capable of capturing multivariate interactions within lipid systems.

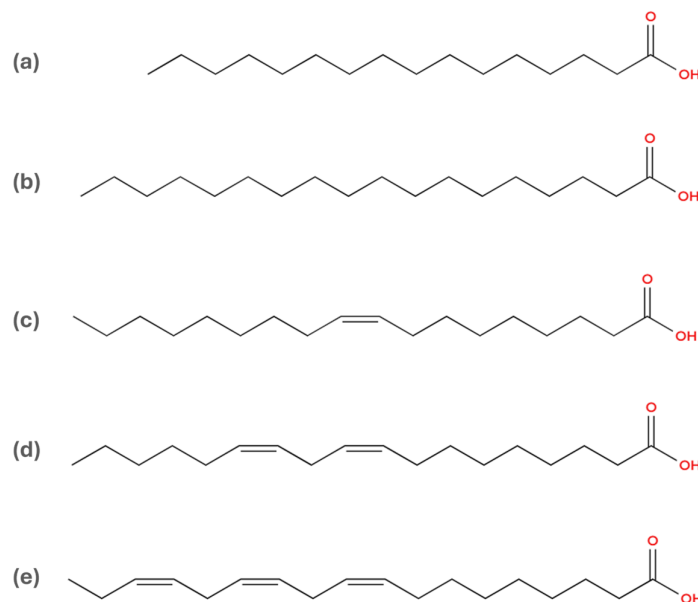


Figure 2. Molecular structure of the main fatty acids covered in this study, generated using MolView (<https://molview.org>), accessed on 14 May 2026 [21]: (a) Palmitic acid; (b) stearic acid; (c) oleic acid; (d) Linoleic acid; (e) Linolenic acid.

In recent years, chemometric techniques have gained prominence as powerful tools for interpreting high-dimensional datasets in fields such as food science, lipidomics, and renewable energy research. Among these techniques, Principal Component Analysis (PCA) stands out for its ability to reduce dimensionality while preserving the variance structure of the data, thereby enabling the identification of latent patterns, correlations, and clustering behaviors. This technique is widely applied in biofuel research as a multivariate statistical tool to reduce complex datasets and identify key variables governing fuel properties, feedstock composition, and process performance. It has been successfully used for feedstock identification and authentication through spectroscopic data, enabling clear discrimination between biodiesel derived from different vegetable oils [22]. Furthermore, PCA has proven effective in monitoring fuel degradation and quality by classifying samples based on physicochemical changes [23]. In addition, it facilitates the understanding of composition–property relationships, such as the influence of fatty acid profiles on viscosity and stability [24,25]. At the process level, PCA helps identify correlations between raw material characteristics and production efficiency, supporting optimization strategies [26].

PCA has been successfully applied to classify oils according to botanical origin, detect adulteration, and explore the influence of compositional variables on functional attributes. Its capacity to integrate multiple physicochemical parameters makes it particularly suitable for investigating the complex relationships between fatty acid profiles and the performance characteristics of biodiesel and biolubricants.

Despite the recognized importance of fatty acid composition in determining the quality of lipid-derived fuels and lubricants, comprehensive multivariate analyses that simultaneously evaluate oxidative stability and viscosity in relation to detailed fatty acid

profiles remain relatively scarce. Most existing studies focus either on isolated correlations, specific subsets of fatty acids, or on compositional classification, which may overlook broader structural patterns that govern product behavior.

Although the fundamental relationships between fatty acid composition, degree of unsaturation, oxidation stability, viscosity, and ester structure are well established in the literature (particularly for biodiesel systems) their systematic integration within a unified multivariate framework remains limited. This is especially true in the case of biolubricants, where fewer studies simultaneously relate feedstock fatty acid profiles to multiple physicochemical properties of the final products.

In this context, although Principal Component Analysis (PCA) has been widely applied in biodiesel research, its use has typically been limited to exploratory classification or to direct correlations between individual variables and physicochemical properties. In contrast, the present work proposes a combined framework integrating PCA-based compositional mapping, clustering-derived regionalization, and variance decomposition. This approach enables the systematic decoupling of fatty acid composition and alcohol-derived structural effects, providing deeper insight into structure–property relationships and supporting more rational strategies for feedstock selection and process optimization.

Consequently, the present work does not aim to demonstrate these relationships for the first time, but rather to organize and interpret them within a coherent multivariate structure. By combining PCA-based compositional mapping, clustering-derived regionalization, and variance decomposition, the proposed approach provides a systematic strategy to disentangle and compare the relative contributions of fatty acid composition and alcohol structure across biodiesel and biolubricant systems. This perspective enables a more integrated understanding of structure–property relationships, particularly in areas where experimental data remain more fragmented, such as in polyol-based biolubricants.

This work proposes a PCA-based multivariate framework to decouple the effects of fatty acid composition and alcohol structure on oxidation stability and viscosity in biodiesel and biolubricants produced from a diverse set of lipid feedstocks. By applying PCA to an integrated dataset comprising detailed fatty acid profiles and key physicochemical properties, this work seeks to uncover multivariate relationships that govern product performance. The findings are expected to contribute to a more comprehensive understanding of how lipid structural features influence the functional attributes of renewable fuels and lubricants, ultimately informing feedstock selection, formulation strategies, and quality control practices within the bio-based materials sector.

2. Materials and Methods

Figure 3 illustrates the conceptual framework underlying the present study, providing a structured representation of how complex compositional and structural effects are disentangled in biodiesel and biolubricant systems. The workflow begins with the fatty acid methyl ester (FAME) composition, which defines the primary chemical space of the system through variables such as chain length and degree of unsaturation. Through Principal Component Analysis (PCA), this high-dimensional compositional information is reduced to a lower-dimensional space (PC1–PC2), where the dominant variability—primarily associated with the degree of unsaturation—is captured. This transformation enables a more tractable interpretation of compositional differences across samples while preserving the essential chemical information. Subsequent clustering of the PCA scores allows the identification of compositional regions that group samples with similar chemical profiles. This regionalization step is critical, as it effectively controls for the dominant source of variability (fatty acid composition), transforming a heterogeneous dataset into subsets of chemically comparable systems.

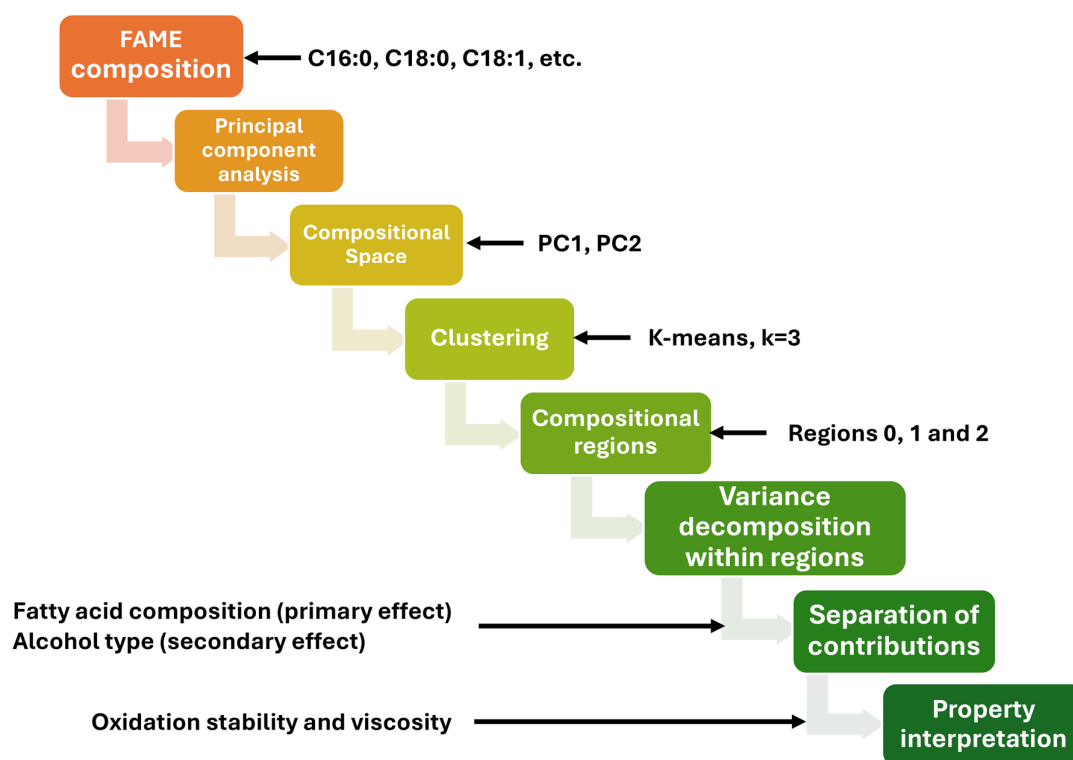


Figure 3. Conceptual framework used to decouple the effects of fatty acid composition and alcohol type on physicochemical properties. The approach combines principal component analysis (PCA) to define the compositional space, clustering to identify homogeneous regions, and variance decomposition within each region to separate primary (fatty acid composition) and secondary (alcohol structure) contributions. This framework enables the interpretation of oxidation stability and viscosity in terms of independent structural effects.

Within each compositional region, variance decomposition is applied to quantify the relative contribution of alcohol type to the observed variability in physicochemical properties. This approach enables the separation of primary effects, driven by fatty acid composition, from secondary effects associated with the alcohol used in transesterification or further processing.

As a result, the framework provides a systematic strategy to decouple and interpret structure–property relationships in complex lipid-derived systems. In particular, it reveals that oxidation stability is largely governed by compositional factors linked to unsaturation, whereas viscosity exhibits a stronger dependence on alcohol structure under controlled compositional conditions. Overall, this integrated methodology transforms a multidimensional and potentially confounded dataset into a sequence of analytically tractable steps, enabling a more rigorous and interpretable assessment of the independent roles of composition and molecular structure in determining final product properties.

2.1. Data Collection and Dataset Construction

A dataset was constructed using both experimental results obtained in this work and complementary data collected from peer-reviewed literature [27–49]. Each observation corresponded to a final fuel or lubricant product derived from vegetable oils or animal fats, including both biodiesel (fatty acid alkyl esters) and biolubricants produced via double transesterification.

The final dataset consisted of 108 samples, including 17 experimental samples generated in this study and 91 samples extracted from the literature. Each observation was considered an independent data point.

The origin of each sample (experimental or literature-derived) is explicitly indicated in the dataset (see Supplementary Table S1). Only studies reporting, at minimum, the fatty acid composition of the feedstock oil and at least one physicochemical property of the resulting product were considered. When multiple operating conditions were reported for the same base product, each formulation was treated as an independent observation.

It should also be noted that the dataset exhibits an inherent imbalance between biodiesel and biolubricant samples. This reflects the current research landscape, where a substantially larger number of studies have systematically explored the relationship between fatty acid composition and physicochemical properties in biodiesel systems. In contrast, comparable studies on biolubricants—particularly those linking feedstock fatty acid profiles with oxidation stability and viscosity—remain more limited. Consequently, the lower representation of biolubricant samples may influence the statistical robustness of certain comparisons and should be considered when interpreting the results.

2.2. Fatty-Acid Composition and Derived Descriptors

For each feedstock oil, the fatty-acid methyl ester (FAME) composition was considered. The five most representative fatty acids (typically accounting for more than 90% of the total composition) were selected for analysis. These generally included palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3) acids, depending on the origin of the oil.

From these data, the following compositional descriptors were calculated:

- Total polyunsaturated fatty acids (Σ PUFA), defined as the sum of fatty acids containing two or more double bonds.
- Degree of unsaturation (DU), calculated as the weighted sum of the number of double bonds in the fatty-acid profile, according to Equation (1):

$$DU = \sum (\%FA_i \times n_i) \quad (1)$$

where $\%FA_i$ is the mass percentage of fatty acid i and n_i is the number of double bonds.

- Long-chain factor (LCF), used as a structural descriptor to account for the contribution of saturated fatty-acid chain length. A truncated formulation was adopted based on the dominant saturated fatty acids present in the selected subset, as shown in Equation (2):

$$LCF = 0.5 \cdot \%C16 : 0 + 1.0 \cdot \%C18 : 0 \quad (2)$$

This formulation provides a comparative measure of the structural contribution of saturated fatty acids and is suitable for datasets dominated by C16 and C18 chains.

2.3. Production Pathways and Process Variables

The products analyzed were classified according to their synthesis route, based on the type of alcohol used:

- Biodiesel, obtained by single transesterification of triglycerides with short-chain alcohols (typically methanol or ethanol).
- Biolubricants, obtained via double transesterification, involving an initial biodiesel production step followed by a second transesterification with a higher alcohol (e.g., 2-ethyl-1-hexanol, pentaerythritol, 2-ethyl-2-hydroxymethyl-1,3-propanediol, etc.).

The type of alcohol used in the transesterification process and the final product category (biodiesel or biolubricant) were recorded as categorical variables and later coded for multivariate analysis (MeOH for methanol, EtOH for ethanol and Polyol for higher alcohols). Categorical variables were dummy-coded prior to analysis. In this work, alcohol

type was categorized into methanol, ethanol, and polyols to reflect the main structural distinction between short-chain alcohols and multifunctional alcohols with higher molecular weight. Although the polyol category encompasses compounds with different structures and numbers of hydroxyl groups, they share common features (like increased molecular size and functionality) that result in significantly higher viscosity and distinct intermolecular interactions compared to simple alcohols.

Within the scope of the present analysis, these shared characteristics justify their grouping into a single category, as they produce a clear separation in the property space. Nevertheless, variability within the polyol group is acknowledged and should be considered when interpreting the results.

2.4. Physicochemical Properties and Oxidation Stability

The following physicochemical properties were collected when available in the literature or experimental tests:

- Oxidation stability, expressed as induction time (hours), typically determined using accelerated oxidation methods.
- Kinematic viscosity, reported at 40 °C.

2.5. Data Pre-Processing

Prior to statistical analysis, all continuous variables were examined for consistency of units and measurement conditions. Although the dataset includes both experimental and literature-derived data, the physicochemical properties considered in this study are typically measured using standardized methods. In particular, oxidation stability is commonly determined using accelerated oxidation tests (e.g., Rancimat method), while kinematic viscosity is measured according to established procedures (e.g., UNE-EN 14214, [50]) at controlled temperatures. The widespread adoption of these standardized methods ensures a reasonable level of comparability across different studies. To enhance consistency, only studies reporting measurements under clearly defined and comparable experimental conditions were included. While some variability related to factors such as purification, residual catalyst, or additive content cannot be entirely eliminated, this is considered in the interpretation of the results and discussed as a limitation of the study.

When necessary, values were converted to ensure uniformity across the dataset. Missing values were handled by case-wise exclusion in multivariate analyses, and outliers were assessed using exploratory plots. The initial dataset consisted of 178 samples, of which 86 were excluded due to incomplete data. The final dataset used for analysis (including literature and experimental) comprised 108 samples.

The excluded samples did not show a systematic bias in terms of alcohol type or compositional characteristics, and their removal was not expected to significantly affect the overall trends observed in the dataset.

For multivariate analysis, all continuous variables were autoscaled (mean-centered and scaled to unit variance) to avoid domination of the models by variables with larger numerical ranges.

2.6. Biodiesel and Biolubricant Production and Characterization

Different vegetable oils (rapeseed, high-oleic safflower, sunflower and cardoon) were used for biodiesel and biolubricant production. The oils were provided by CICYTEX (Centro de Investigaciones Científicas y Tecnológicas de Extremadura) in 2026. For biodiesel production, the vegetable oils were subjected to transesterification to produce fatty acid methyl or ethyl esters (FAMES and FAEs) and glycerol. The reaction was carried out using methanol or ethanol in the presence of potassium hydroxide (KOH, 1% *w/w*) as a catalyst.

Since transesterification is a reversible reaction, operating parameters such as temperature (65 and 70 °C for methanol and ethanol, respectively), catalyst concentration (1%), and alcohol-to-oil molar ratio (6:1) were selected to maximize the FAME yield according to previous studies [51]. The experimental setup is included in Figure 4.

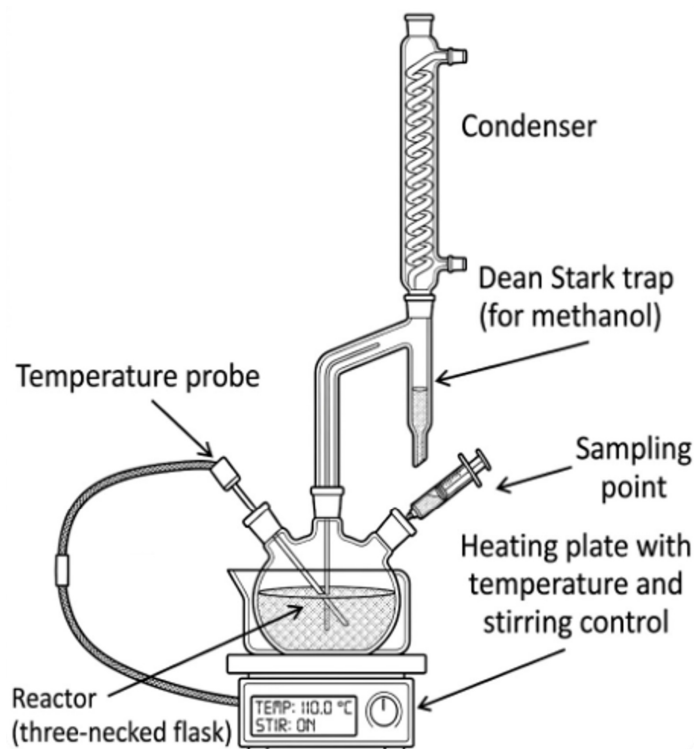


Figure 4. Experimental setup for biodiesel production.

The vegetable oils used for experimental biodiesel and biolubricant production are included in Table 1, with their main compositional characteristics.

Table 1. Vegetable oils used in this study for biodiesel and biolubricant production and their fatty acid profile.

Vegetable Oil	C16:0	C18:0	C18:1	C18:2	C18:3
Rapeseed	4.86	4.15	64.32	17.3	1.12
High-oleic safflower	4.81	2.32	81.05	8.08	0.5
Sunflower	5.42	2.81	28.32	56.02	0.2
Cardoon	11.36	4.3	29.1	50.8	0.6

The process was conducted in a 1 L three-necked reactor. The reaction temperature was selected based on the boiling point of the alcohol used (methanol or ethanol). To prevent alcohol evaporation, a condenser was attached to the reactor. Throughout the reaction, temperature was continuously monitored and controlled. After completion of the reaction, the FAME phase was purified by decantation to remove the glycerol, followed by washing with distilled water. The final product was dried by heating at 100 °C and subsequently stored in amber glass bottles. All experiments were performed immediately after purification and drying. Equally, vegetable oils underwent double transesterification to produce biolubricants (first transesterification with methanol, as explained in previous paragraphs and second transesterification with different superior alcohols). This second reaction was carried out using different alcohols (namely pentaerythritol and 2-ethyl-2-

hydroxymethyl-1,3-propanediol) together with the previously obtained FAMES and sodium methoxide as a catalyst.

The experimental setup was similar to that used for FAME production (see Figure 5), with two main differences: the inclusion of a Dean–Stark trap and a sampling port. The Dean–Stark trap enabled the continuous collection and removal of the methanol generated during the second transesterification reaction, thereby shifting the equilibrium toward biolubricant formation and improving product yield. The sampling point allowed periodic collection of reaction samples to monitor the evolution of residual FAME content by gas chromatography.

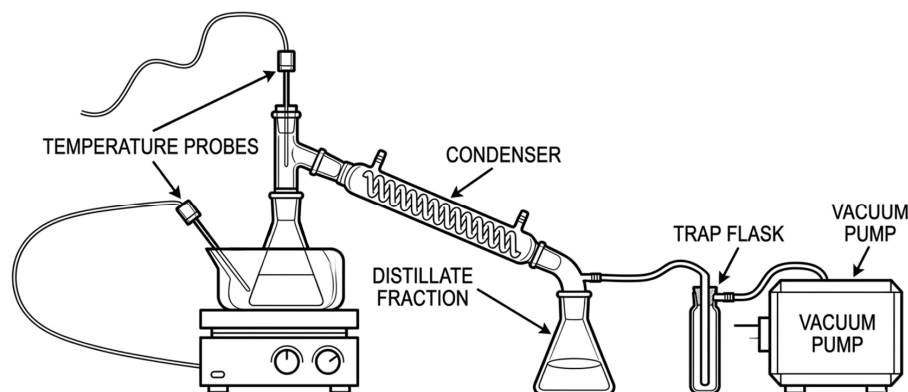


Figure 5. Experimental setup for biolubricant production.

In the second transesterification step, process variables such as reaction temperature, catalyst concentration, alcohol-to-FAME molar ratio, and alcohol type were investigated to optimize biolubricant yield. The reaction conditions are summarized in Table 2, according to previous studies [49,51]. The reaction time was 120 min, with a stirring rate of 500 rpm.

Table 2. Main parameters for biolubricant production (second transesterification).

Alcohol	Alcohol/FAME Molar Ratio	Pressure, mmHg	Temperature, °C	[Catalyst] % w/w
Pentaerythritol	1/3	260	160	1.0
2-ethyl-2-hydroxymethyl-1,3-propanediol	1	400	140	1.0

The reactions were carried out at temperatures close to the boiling point of each alcohol, allowing operation at the highest possible temperature without causing alcohol evaporation, thus maintaining it within the reaction medium. After completion of the reaction, excess alcohol (methanol) was removed by vacuum distillation. For purification of the final product, vacuum filtration was employed to remove residual unreacted alcohol in solid form. This procedure yielded conversion values ranging from 90 to 94% in all cases. The residual fatty acid methyl ester (FAME) content was subsequently quantified by gas chromatography, as described in the following section.

The characterization of fatty acid methyl esters (FAMES) was carried out following the corresponding standards [50]. Gas chromatographic analyses were performed using a Varian 3900 gas chromatograph equipped with a flame ionization detector (FID). Separation was achieved using a Zebron ZB-Wax Plus capillary column (Phenomenex, Torrance, CA, USA), (30 m length, 0.32 mm internal diameter, 0.25 µm film thickness), with a maximum operating temperature of 260 °C.

For each FAME analyzed (in this case, methyl oleate, methyl linoleate, methyl palmitate, methyl linolenate and methyl stearate) individual calibration curves were prepared

using the corresponding analytical standards (Sigma-Aldrich, Burlington, MA, USA). Quantification was performed using methyl heptadecanoate (Sigma-Aldrich, Burlington, MA, USA) as an internal standard. All gases used during the chromatographic analyses were supplied by Linde (Munich, Germany), with hydrogen and synthetic air employed as fuel and oxidant gases, respectively. A known amount of the internal standard was added to each sample prior to analysis. Calibration curves were constructed for each FAME using analytical standards, allowing the determination of response factors relative to the internal standard. Concentrations were calculated from the ratio of peak areas (FAME/internal standard), correcting for differences in detector response. Although flame ionization detection (FID) does not provide absolute concentration measurements, its response is proportional to the mass of carbon, enabling accurate quantification when combined with calibration and internal standardization.

Additionally, physicochemical properties were determined according to standardized methods. Kinematic viscosity was determined at 40 °C following ISO 3104:1994, using an Ostwald viscometer [52].

Oxidative stability was determined using the Rancimat method [53]. Approximately 3 g of each sample was placed in a test tube, through which synthetic air (10 L h⁻¹, Linde, UK) was continuously bubbled while the tube was maintained at 100 °C. The volatile oxidation products generated during the experiment were carried by the airflow into a vessel containing 50 mL of deionized water. As oxidation progressed, secondary oxidation products dissolved in the water, leading to an increase in conductivity.

Conductivity was continuously monitored using a conductivity meter (Crison EC-Meter GLP31+, Barcelona, Spain). The conductivity profile exhibits two distinct stages: an initial phase characterized by a nearly constant conductivity with a flat slope, followed by a second phase in which a sharp increase in conductivity occurs due to the rapid formation and release of auto-oxidation by-products. The change in slope was assigned to oxidation stability or induction period, expressed in hours.

All experimental measurements, including kinematic viscosity and oxidation stability (Rancimat), were performed in triplicate. The values reported correspond to the mean of the measurements, while the associated variability was evaluated in terms of standard deviation.

2.7. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) was performed to identify the main sources of variability in the dataset based on fatty acid composition. The analysis included both individual fatty acid variables (C16:0, C18:0, C18:1, C18:2, C18:3) and derived indices (polyunsaturated content, total number of double bonds, and structural index) to capture both detailed and global compositional features. The properties considered were kinematic viscosity (mm²·s⁻¹) and oxidation stability (h), which were evaluated separately to assess their dependence on composition and alcohol type.

In addition to individual fatty acid contents, derived indices such as PUFA, degree of unsaturation (DU), and long-chain factor (LCF) were included to facilitate interpretation of the compositional space. It is acknowledged that these variables are not fully independent, as they are calculated from the same underlying fatty acid data.

However, their inclusion provides a more direct representation of chemically meaningful features, such as overall unsaturation and chain structure, which are closely related to physicochemical properties. PCA inherently accounts for correlations among variables, allowing redundant information to be incorporated into the same principal components without compromising the overall analysis.

PCA was conducted in JASP 0.97.0 using the correlation matrix to ensure proper scaling of variables, and no rotation was applied to preserve the original orthogonal structure of the components. The selection of principal components was initially based on the Kaiser criterion (eigenvalues greater than one). In addition, the scree plot (Figure 6) was used to identify a clear inflection point in the explained variance, supporting the retention of the first principal components. Furthermore, the cumulative explained variance and the stability of the loading patterns were considered to ensure a robust and interpretable representation of the dataset. The use of multiple complementary criteria for principal component selection is recommended to avoid over- or underestimation of the number of relevant components and to improve the reliability of PCA interpretation, as discussed in the literature [54]. The eigenvalues and explained variance of the principal components are provided in Table S2. According to the Kaiser criterion (eigenvalues > 1), two principal components were retained, explaining 77.7% of the total variance.

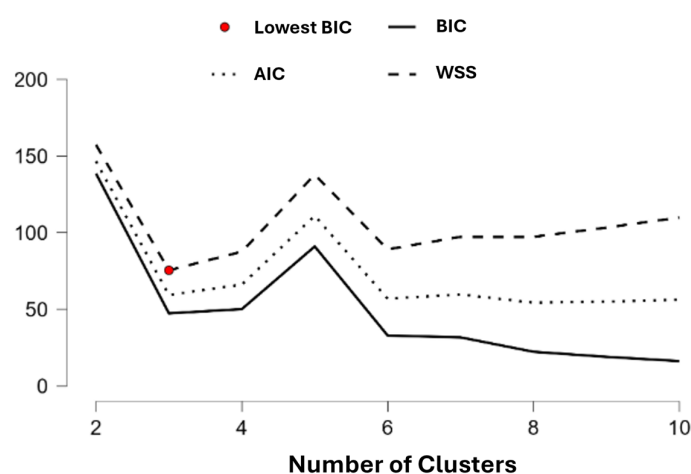


Figure 6. Determination of the optimal number of clusters using the elbow method applied to PCA scores (PC1–PC2), where AIC is Akaike Information Criterion, BIC is Bayesian Information Criterion and WSS is Within-Cluster Sum of Squares.

All variables were standardized prior to PCA to eliminate scale effects and ensure that differences in magnitude (e.g., viscosity ranges between biodiesel and polyol-derived products) do not dominate the analysis. Although biodiesel and biolubricants differ significantly in their physicochemical properties, both systems were analyzed within a unified framework to explore the combined influence of fatty acid composition and alcohol structure. In this context, the broader variability observed for certain properties (such as viscosity) reflects meaningful structural differences rather than analytical bias.

These combined criteria ensure a robust and justified selection of principal components. The quality of the PCA model was evaluated through the explained variance ratio, with the first two principal components used to represent the dataset in a reduced dimensional space. All variables included in the PCA were analyzed using the correlation matrix to ensure appropriate scaling and comparability. Derived indices were incorporated to enhance the representation of structural features related to unsaturation. The robustness of the PCA model was assessed by evaluating stability of loadings and explained variance. Cluster analysis was performed using k-means with $k = 3$, selected based on the elbow method (WSS) and visual inspection of the PCA score distribution. The algorithm was initialized with random centroids and repeated 100 times to ensure stability.

The optimal number of clusters was determined using a combination of the elbow method and silhouette analysis applied to the PCA scores (PC1 and PC2). The elbow method, based on the within-cluster sum of squares (WCSS), showed a sharp decrease

from $k = 1$ (591.5) to $k = 3$ (136.0), followed by a more gradual reduction for higher values (e.g., $k = 4$: 96.2; $k = 5$: 66.3), indicating a clear inflection point at $k = 3$ (Figure 6).

Silhouette analysis further confirmed the robustness of this solution, with an average coefficient of 0.54 for $k = 3$. Individual cluster values (0.49, 0.49, and 0.66) indicate good internal cohesion and clear separation between groups.

Increasing the number of clusters beyond $k = 3$ did not significantly improve compactness and resulted in lower silhouette values and reduced interpretability.

Based on the agreement between the elbow method, silhouette analysis, and chemical interpretability, $k = 3$ was selected as the optimal number of clusters, providing a robust and objective basis for cluster selection.

The robustness of the clustering solution was further supported by the consistent sample distribution within regions and their chemical interpretability.

The correlation matrix of the variables included in the PCA is provided in the Supplementary Materials (Table S3). As expected, strong correlations are observed between individual fatty acids and derived descriptors (e.g., C18:2–polyunsaturated, PUFA–double bonds, and structural index–C18:1), indicating partial redundancy among variables. This multicollinearity supports the use of PCA as a dimensionality reduction technique, allowing the extraction of orthogonal components that capture the dominant variance structure of the dataset.

2.8. Definition of Compositional Regions

To control the influence of fatty acid composition, the PCA score space was partitioned into compositional regions. These regions were defined based on clustering of the PCA scores, resulting in subsets of samples with comparable chemical characteristics. This regionalization allowed the dataset to be divided into chemically homogeneous groups, minimizing variability associated with fatty acid composition. The suitability of the dataset for PCA was assessed using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett’s test of sphericity. The KMO value (0.78) indicates that the correlation structure of the dataset is appropriate for dimensionality reduction. In addition, Bartlett’s test was statistically significant ($p < 0.001$), confirming that the variables are sufficiently correlated to justify the application of PCA.

2.9. Assessment of Alcohol Effect

Within each compositional region, the effect of alcohol type on physicochemical properties was evaluated. Viscosity and oxidation stability were analyzed as response variables. To isolate the contribution of alcohol independently of fatty acid composition, the variability within each region was decomposed into total variance and variance between alcohol groups. The relative contribution of alcohol was quantified using the Fraction Alcohol Effect (Equation (3)), defined as:

The Fraction Alcohol Effect (FAE) was defined as a variance-based metric equivalent to the η^2 (eta-squared) effect size used in one-way ANOVA. It quantifies the proportion of total variance in a given property that is attributable to differences between alcohol groups within each compositional region:

$$\text{FAE} = \frac{SS_{\text{between}}}{SS_{\text{total}}} \quad (3)$$

where SS_{between} represents the sum of squares between alcohol groups and SS_{total} corresponds to the total sum of squares within the region (Equations (4) and (5)).

$$SS_{\text{between}} = \sum n_j (\mu_j - \mu)^2 \quad (4)$$

$$SS_{total} = \sum (x_i - \mu)^2 \quad (5)$$

where n_j is the number of observations in alcohol group j , μ_j is the mean value of that group, μ is the overall mean within the region, and x_i represents individual observations.

FAE values close to 0 indicate that variability is primarily governed by fatty acid composition, whereas values approaching 1 indicate a dominant contribution of alcohol type.

Although FAE is mathematically equivalent to η^2 , it is used here as a descriptive metric to assess the relative importance of alcohol as a source of variability, rather than as a formal inferential effect size estimator. Given the heterogeneous and partially unbalanced nature of the dataset (particularly in regions with limited representation of certain alcohol types) the FAE values should be interpreted as indicative of trends rather than statistically rigorous measures of effect magnitude.

Therefore, FAE is employed as an exploratory tool to compare the relative influence of alcohol across compositional regions, and its interpretation is supported by complementary statistical analyses. In this context, the statistical significance of the alcohol effect was evaluated using one-way ANOVA, with differences considered significant at $p < 0.05$.

3. Results and Discussion

Although biodiesel and polyol-derived biolubricants differ significantly in their physicochemical properties, both systems are analyzed within a common multivariate framework to explore underlying structural relationships. Data standardization and the use of PCA ensure that differences in scale do not dominate the analysis, while the definition of compositional regions enables comparisons within chemically comparable subsets.

Principal Component Analysis (PCA) was performed using both individual fatty acid variables and derived indices (polyunsaturated content, total double bonds, and structural index). The PCA results describe the compositional structure of the dataset, particularly in terms of fatty acid profiles and derived descriptors. As observed in Table 3, the first principal component (PC1) explains 52.7% of the total variance and is strongly associated with the global degree of unsaturation, as evidenced by the high positive loadings of polyunsaturated fatty acids and total double bonds. Negative contributions from monounsaturated species and structural parameters further indicate that PC1 represents a compositional gradient from saturated/monounsaturated to highly unsaturated profiles. The second principal component (PC2), explaining an additional 25.0% of the variance, captures secondary compositional differences among fatty acid types.

Table 3. Principal component analysis results: model fit, loadings, and component characteristics.

Model Fit (Chi-Squared Test)	Value	df	<i>p</i>
Model	6412.203	13	<0.001
Component loadings	PC1	PC2	Uniqueness
Polyunsaturated	0.996		0.008
C18:2	0.948		0.095
Structural index	−0.871	0.448	0.041
Double bonds in total	0.859	0.496	0.017
C18:1	−0.777	0.612	0.022
C16:0		−0.724	0.376
C18:0		−0.720	0.411
C18:3			0.816
Component characteristics	Eigenvalue	Proportion var.	Cumulative
Component 1	4.216	0.527	0.527
Component 2	1.998	0.250	0.777

Together, the first two components account for 77.7% of the total variance, indicating a well-structured dataset in which the primary chemical variability is dominated by the degree of unsaturation. The observed grouping of samples in the PCA space provides a compositional framework that can be related, in a subsequent step, to physicochemical properties. Thus, while PCA does not directly establish relationships between properties such as viscosity and oxidation stability, the distribution of samples according to their composition enables an indirect interpretation of how these properties vary across different compositional regions.

Figure 7 shows that PC1 captures the largest share of variance, indicating that a single dominant underlying factor governs most of the variability in the dataset. The second principal component (PC2) also contributes a significant portion of the variance, while subsequent components explain only minor additional variability. Together, the first two principal components capture most of total variance, demonstrating that the multidimensional dataset can be effectively reduced to a two-dimensional representation without substantial loss of information.

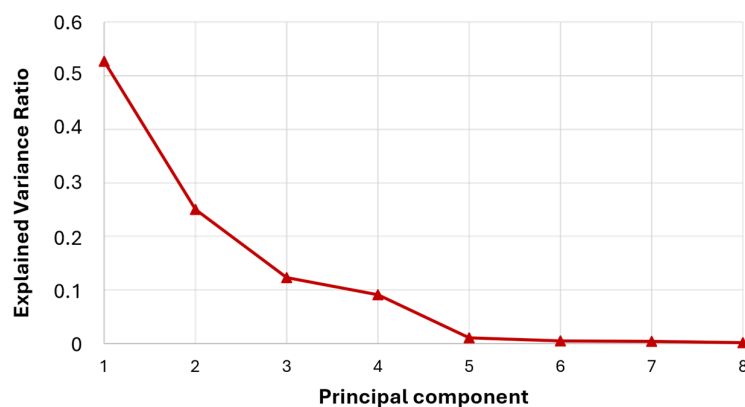


Figure 7. Explained variance of principal components obtained from PCA. The first principal component (PC1) accounts for the largest proportion of variance (0.53), followed by PC2 (0.25) and PC3 (0.12). Subsequent components contribute progressively smaller amounts of variance, with PC4–PC6 each explaining less than 0.20, and components beyond PC5 contributing marginally (≤ 0.02). This distribution indicates that most of the variability in the dataset is captured by the first two to three principal components.

The dominance of PC1 suggests that a primary physicochemical factor drives most of the variation. Based on the loading analysis, this factor corresponds to the degree of unsaturation, as variables related to double bonds and polyunsaturated fatty acids strongly contribute to this component. PC2 captures secondary structural differences in fatty acid composition, such as the balance between saturated and monounsaturated species. This indicates that the dataset is highly structured, with most of the variability being governed by a limited number of chemically meaningful factors, primarily the degree of unsaturation. Additionally, the fact that only two or three components are required to explain most of the variance validates the use of PCA as an effective dimensionality reduction and exploratory tool for this system. The scree plot (Figure 7) shows a clear inflection point after the first principal components, indicating a natural cutoff in the explained variance. Furthermore, the cumulative variance associated with the first components exceeds 77%, confirming that most of the variability in the dataset is captured. Finally, the retained components exhibit stable and chemically meaningful loading patterns, particularly in relation to the degree of unsaturation and fatty acid composition. The combination of these criteria supports the selection of the principal components used in this study, rather than relying on a single selection rule.

It should be noted that defining a single fixed oil is inherently challenging, as even nominally identical feedstocks exhibit natural variability in fatty acid composition. Consequently, attempts to isolate the effect of alcohol by selecting a single oil may still introduce compositional variability, potentially confounding the interpretation. To address this limitation, the analysis was performed considering the full dataset, while interpreting trends along the principal compositional axis (PC1) and evaluating alcohol-dependent deviations within comparable compositional regions. This approach allows disentangling the influence of fatty acid composition and alcohol type without artificially constraining the dataset.

Therefore, Figure 8 shows the PCA score plot. To facilitate interpretation, the PCA score space was further partitioned into three compositional regions using an unsupervised clustering approach (k-means). The PCA score plot reveals a well-defined chemical organization of the samples, in which the first principal component (PC1) captures the dominant compositional gradient associated with the degree of unsaturation of fatty acids. Samples are distributed along this axis from saturated profiles to increasingly unsaturated compositions. The second principal component (PC2) captures secondary variations related to more subtle compositional differences, such as the relative contribution of mono- versus polyunsaturated species.

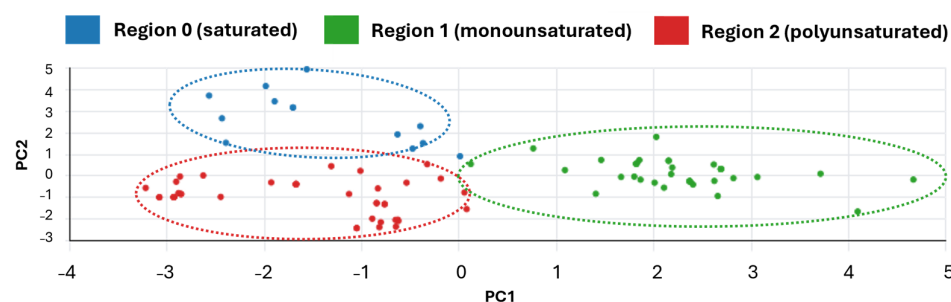


Figure 8. PCA score plot (PC1 vs. PC2). The data are enclosed by dashed circles (blue for Region 0, green for Region 1 and red for Region 2).

The application of k-means clustering to the PCA scores resulted in the identification of three distinct regions within the compositional space. In this way, the number of clusters was selected based on visual inspection of the PCA score distribution and supported by within-cluster variance minimization criteria. These regions correspond to:

- Region 0: predominantly saturated or low-unsaturation profiles.
- Region 1: intermediate compositions, typically enriched in monounsaturated fatty acids.
- Region 2: highly unsaturated profiles, characterized by elevated levels of polyunsaturated fatty acids.

Importantly, these regions are defined exclusively based on fatty acid composition, without considering alcohol type or physicochemical properties. The partitioning of the PCA space into compositional regions provides a robust framework to control the effect of fatty acid composition when analyzing physicochemical properties. Within each region, samples exhibit comparable chemical characteristics, allowing systematic deviations to be attributed to other factors. This approach effectively transforms the dataset into a series of chemically homogeneous subsets, enabling the isolation of secondary effects that would otherwise be masked in the global analysis.

By reducing compositional variability within each region, it becomes possible to evaluate the influence of the alcohol used in esterification independently of fatty acid composition. In this context, variations observed along PC1 are primarily governed by fatty acid composition and are therefore associated with intrinsic chemical properties such

as oxidation stability. Also, deviations observed within each region, not aligned with PC1, reflect additional structural effects, notably those introduced by the alcohol moiety. As a result, the combination of PCA and clustering provides a powerful strategy to disentangle compositional and structural effects in multivariate datasets, revealing that while fatty acid composition defines the primary chemical framework, alcohol type introduces secondary variations that can be systematically analyzed within comparable compositional regions.

Following the PCA by regions, Table 4 presents the variance decomposition of oxidative stability and viscosity, showing the relative contribution of alcohol type to total variability across different regions and properties. It provides a quantitative complement by explicitly partitioning the variance and isolating the contribution of alcohol type. While PCA identifies dominant patterns and structural relationships within each region, it does not quantify the relative importance of specific factors. This table addresses that gap by showing how much of the variability in oxidative stability and viscosity can be attributed to alcohol differences, thereby clarifying whether alcohol type acts as a primary or secondary driver under each regional condition.

Table 4. Variance decomposition of alcohol fraction effects on oxidative stability (OS) and viscosity by region (Region 1 was excluded due to lack of data), where n.d. = not determined.

Region	0	1	2
Total <i>n</i>	37	19	36
Methanol	32	17	28
Ethanol	2	2	2
Polyol	3	0	6

Property	OS	Viscosity	OS	Viscosity	OS	Viscosity
Total variance	19.02	0.118	n.d.	n.d.	13.34	791.34
Variance between alcohols	12.65	0.0649	n.d.	n.d.	1.38	99.73
FAE	0.665	0.549	n.d.	n.d.	0.104	0.126

The dataset exhibits an unbalanced distribution of samples among alcohol types, with a predominance of biodiesel samples and a lower representation of polyol-derived biolubricants in some compositional regions. This imbalance may affect the statistical power and robustness of the analysis, particularly in regions where the number of polyol samples is limited. In such cases, variance estimates and effect size metrics (e.g., Fraction Alcohol Effect, FAE) may be more sensitive to sample variability. Therefore, results in regions with fewer observations should be interpreted with caution, especially when evaluating the variance decomposition presented in Table 4.

The distribution of samples across compositional regions and alcohol types was evaluated to ensure meaningful comparisons. Region 0 ($n = 37$) and Region 2 ($n = 36$) include all alcohol types (methanol, ethanol, and polyols), enabling a direct assessment of alcohol effects within these compositional spaces. Region 1 contains 19 samples, but only biodiesel-derived systems (methanol and ethanol), with no polyol-derived samples. This imbalance should be considered when interpreting the alcohol effect in this region. Overall, the dataset provides sufficient representation for comparative analysis, although variability in alcohol distribution across regions may influence the robustness of specific comparisons.

First, Region 0 shows a strong dependence on alcohol type, particularly for oxidative stability. The total variance for oxidative stability is 19.02, of which 12.65 is explained by differences between alcohols, resulting in a fraction alcohol effect of 0.665. This means that approximately 66.5% of the variability in oxidative stability is driven by the choice of alcohol, which represents a substantial proportion. In practical terms, this implies that

in Region 0, selecting between MeOH, EtOH, or polyols has a dominant influence on oxidation resistance, likely overriding, to some extent, differences in the oil feedstock itself. For viscosity in Region 0, the fraction alcohol effect is also notable (0.549), meaning that more than half of the observed variability (54.9%) is explained by alcohol type. Although the absolute variance is small (0.118), the relative contribution of alcohol remains significant. This reinforces the idea that, even when viscosity differences are less pronounced in magnitude, alcohol chemistry still plays a key structural role in determining fluid behavior.

In contrast, Region 2 exhibits a weaker influence of alcohol fraction for both properties. The Fraction Alcohol Effect for oxidative stability decreases to 0.104, indicating that only 10.4% of the variability is explained by alcohol differences. Similarly, for viscosity, despite a high total variance (791.34), only a limited fraction is attributable to alcohol (99.73), resulting in a Fraction Alcohol Effect of 0.126. These results suggest that, in Region 2, other factors (most likely related to the intrinsic composition of the oils, such as the degree of unsaturation or structural index) play a more dominant role than alcohol type. The large variability in viscosity can be explained by the coexistence of biodiesel samples and polyol-derived biolubricants within this region. While biodiesel exhibits relatively low viscosity values within a narrow range, polyol-based products show significantly higher viscosity, leading to a broad dispersion of values and increased total variance.

However, the magnitude of the total variance suggests that other factors, particularly feedstock-related variables, dominate the variability in viscosity. This suggests that, although polyols exert a strong intrinsic effect on viscosity, their relative contribution becomes less significant when larger sources of variation are present.

Within each compositional region, viscosity and oxidation stability were averaged for each alcohol type and compared across groups. Figure 9 shows the main results concerning this point, for the four oils used in this study and the different alcohols added for biodiesel and biolubricant production (grouped into MeOH and EtOH for biodiesel and Polyol for biolubricant production).

The comparative analysis of the different vegetable oils and alcohols presented in the figure reveals several clear relationships between the structural properties of the feedstocks and their resulting oxidation stability. As observed in Figure 9a, sunflower and cardoon oils exhibit the highest total number of double bonds (approximately 130–140), along with significant polyunsaturation. In contrast, high-oleic safflower and rapeseed oils display fewer double bonds and lower polyunsaturated content, with a comparatively higher structural index in the case of high-oleic safflower. These compositional differences (which are similar to other results found in the literature [19,55]) are crucial because a higher number of double bonds, particularly in polyunsaturated fatty acids, is directly associated with increased susceptibility to oxidative degradation. Therefore, oils such as sunflower and cardoon are expected to be less stable than high-oleic safflower oil, which has a more favorable, less unsaturated structure [42,47,49].

This structural trend is consistent with the oxidation stability results shown in Figure 9b, where the alcohol type also plays a decisive role. Transesterification with methanol and ethanol to produce biodiesel exhibits similar oxidation stability values, both around 5 h, indicating relatively good resistance to oxidative degradation regardless of slight differences in polarity or molecular size. In contrast, polyol-based biolubricants show significantly lower oxidation stability, around 3 h, indicating that the presence of multiple hydroxyl groups may influence oxidation mechanisms and reduce oxidative robustness, as previously reported for polyol ester systems [47,56]. This suggests that while polyols may confer other desirable properties (such as higher viscosity or improved lubrication), they could compromise oxidative stability, requiring the addition of antioxidants. This behavior is also consistent with the literature, where bio-based lubricants often exhibit

lower oxidation stability compared to conventional systems [57], requiring formulation strategies such as antioxidant addition to improve performance [58].

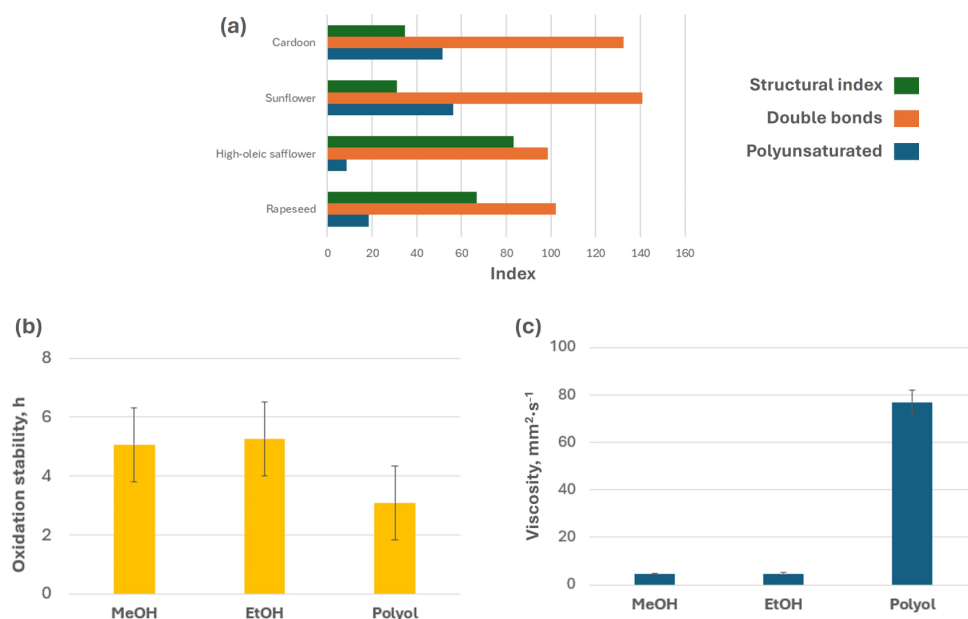


Figure 9. Main results derived from biodiesel and biolubricants obtained experimentally: (a) data based on fatty acid composition of original vegetable oils; (b) average oxidation stability according to the alcohol used; (c) average viscosity according to the alcohol used.

The viscosity results observed in Figure 9c further complement this interpretation. Polyol-based biolubricants showed higher viscosity (close to 80 mm²·s⁻¹) compared to biodiesel obtained from transesterification with methanol and ethanol (both around 4–5 mm²·s⁻¹, in compliance with the standard [50]). This trend suggests that increasing molecular complexity and functionality is associated with the formation of more viscous and resistant fluids, although a reduction in oxidation stability is also observed. In practical terms, this indicates a potential trade-off between rheological properties and chemical stability: more complex alcohols tend to increase viscosity, while their influence on oxidative behavior appears less favorable. In line with this observation, polyol-derived samples tend to exhibit higher viscosity values compared to biodiesel systems, as confirmed by statistical analysis ($p < 0.05$), although variability within each group remains relevant. This behavior is consistent with the larger molecular size and multifunctional structure of polyol esters, which promote stronger intermolecular interactions. Nevertheless, these relationships should be interpreted considering the variability within each group and the limitations of the dataset.

It should be noted that biodiesel and polyol-derived biolubricants are designed for different applications and therefore follow different performance criteria. For biodiesel, key properties such as viscosity and oxidation stability are regulated by standards such as EN 14214 [50], which specifies a viscosity range of 3.5–5.0 mm²·s⁻¹ at 40 °C and a minimum oxidation stability of 8 h.

In contrast, polyol-derived biolubricants are not subject to the same specifications and are typically characterized by significantly higher viscosity values due to their larger molecular structure. Therefore, the broader range of values observed in this study reflects not only compositional effects, but also the inherently different functional requirements of both types of products.

Overall, these results suggest that both fatty acid composition of the oil and the type of alcohol used in its derivatization or processing critically influence final performance.

Oils with lower unsaturation (such as high-oleic safflower) are inherently more stable, while simpler alcohols (methanol and ethanol) favor higher oxidation stability. Conversely, highly unsaturated oils combined with polyols are likely to produce materials that are more viscous but significantly less resistant to oxidation. These findings highlight the importance of carefully selecting both feedstock and alcohol depending on the intended application, especially when balancing stability and physical properties is essential.

From a molecular perspective, biodiesel consists of fatty acid alkyl esters (FAAE), in which the hydrocarbon chain derives from the original triglyceride and the alkyl group from the alcohol. Both fragments contribute to intermolecular interactions and ultimately determine macroscopic properties such as viscosity, density, cold flow behavior, and oxidative stability [54,59,60].

Fatty acid composition primarily governs molecular packing, flexibility, and chemical reactivity. Saturated fatty acids promote stronger van der Waals interactions and tighter packing, leading to higher viscosity and enhanced oxidative stability. In contrast, unsaturated fatty acids introduce double bonds that disrupt packing, reducing viscosity and improving low-temperature behavior, but increasing susceptibility to oxidation due to their reactive sites [30,31]. As transesterification does not alter the fatty acid backbone, the oxidative stability of biodiesel and biolubricants is largely inherited from the original feedstock. Thus, oils rich in polyunsaturated fatty acids tend to produce fuels with lower stability, while saturated and monounsaturated profiles enhance resistance [59,60].

In parallel, the alcohol used in transesterification modifies the size and steric environment of the ester group. Short-chain alcohols (e.g., methanol, ethanol) produce esters with lower molecular weight, weaker intermolecular interactions, and lower viscosity, but higher susceptibility to oxidation [61]. Increasing the alcohol chain length results in bulkier and more hydrophobic esters, enhancing molecular interactions and viscosity, while providing partial steric protection that may improve oxidative resistance. Branching further influences the balance between viscosity and stability.

This progressive structural modification explains the transition from fuel-like to lubricant-like behavior, with short-chain esters acting as low-viscosity fuels and long-chain esters exhibiting the higher viscosity and film-forming capacity required for lubrication. Consequently, oxidative stability becomes critical in biolubricants, often requiring feedstocks with low polyunsaturation or the use of additives and chemical modifications to enhance durability [19].

Overall, biodiesel and biolubricants can be understood as a continuum defined by molecular architecture. By tuning both fatty acid composition and alcohol structure, it is possible to tailor viscosity, stability, and overall performance, highlighting transesterification as a versatile tool for designing bio-based fluids.

However, several limitations should be considered in this study. First, the dataset combines experimental results with literature data, which may introduce variability due to differences in analytical methods and operating conditions. Additionally, the distribution of samples across compositional regions and alcohol types is not fully balanced, which may affect the robustness of the variance decomposition and the estimation of the alcohol effect. In particular, the presence of high-viscosity polyol-derived samples introduces potential outliers that lead to a highly dispersed and partially multimodal distribution of viscosity values, which can significantly influence statistical descriptors such as variance. Furthermore, the use of PCA, as a linear method, may not fully capture complex non-linear relationships between fatty acid composition and physicochemical properties, especially in systems where interactions between variables are not purely additive. The analysis was restricted to oxidation stability and viscosity, without considering other relevant performance parameters such as cold-flow behavior or long-term stability. Finally, although efforts were

made to standardize units and conditions, variability associated with different experimental protocols (e.g., oxidation stability methods, temperature conditions) may introduce systematic bias. Alternative approaches such as t-distributed stochastic neighbor embedding (t-SNE), uniform manifold approximation and projection (UMAP), and kernel PCA have been proposed to better represent nonlinear structures in multivariate data [62,63]. However, PCA was selected in this work due to its interpretability and its ability to provide direct insight into the contribution of original variables to the principal components. In addition, the consistency of the observed trends with well-established structure–property relationships suggests that the main variability in the dataset is effectively captured within a linear framework. Further exploration using nonlinear methods could provide additional insight into subtle relationships but is beyond the scope of the present study.

4. Conclusions

This study demonstrates that the physicochemical behavior of biodiesel and biolubricant systems can be systematically interpreted by decoupling the effects of fatty acid composition and alcohol structure using a combined multivariate and statistical approach.

Principal Component Analysis (PCA) indicates that dataset variability is largely explained by a limited number of chemically meaningful factors, primarily associated with the degree of unsaturation. The partitioning of the PCA score space into compositional regions enables the control of fatty acid variability and allows a more rigorous assessment of secondary effects, particularly the role of alcohol type.

Within these regions, viscosity and oxidation stability exhibit distinct behaviors. Viscosity shows a consistent dependence on alcohol structure, with polyol-derived systems displaying significantly higher values ($p < 0.05$). This behavior is consistent with their larger molecular size and stronger intermolecular interactions.

In contrast, the influence of alcohol type on oxidation stability depends on the compositional region. In some cases, only limited variability across alcohols is observed. However, a significant contribution of alcohol type is detected in specific regions (e.g., Region 0), indicating that its effect is modulated by the underlying fatty acid composition.

The introduction of the Fraction Alcohol Effect provides a quantitative framework to evaluate these contributions. This analysis shows that viscosity is strongly influenced by alcohol-derived structural effects. In contrast, oxidation stability is primarily influenced by fatty acid composition in highly unsaturated systems. Nevertheless, alcohol type can play a dominant role in specific compositional regions, highlighting the context-dependent nature of these effects.

Despite the robustness of the proposed framework, some limitations should be acknowledged. These include the combination of experimental and literature data, potential variability due to external factors, and the restriction of the analysis to viscosity and oxidation stability.

Overall, this study provides a consistent framework in which fatty acid composition defines the overall chemical space. Alcohol type introduces a secondary (yet in some cases dominant) modulation of physicochemical properties depending on the compositional region. This approach offers useful insights for the rational design of biodiesel and biolubricant formulations.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app16126049/s1>, Table S1: Characteristics of the Study Dataset; Table S2: Eigenvalues and explained variance of principal components; Table S3: Correlation matrix of variables used in PCA.

Funding: This research was funded by the European Union, European Regional Development Fund and Regional Government of Extremadura (GR24127).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The author would like to thank the support provided by his family during his scientific career. Thank you so much for your endless love. The author would also like to thank the European Union, European Regional Development Fund and Regional Government of Extremadura (Este trabajo ha sido cofinanciado al 85% por la Unión Europea, Fondo Europeo de Desarrollo Regional, y la Junta de Extremadura. Autoridad de Gestión. Ministerio de Hacienda (GR24127)). During the preparation of this manuscript/study, the author used M365 Copilot, based on the GPT-5 Chat model from OpenAI, for the purposes of reviewing the writing of some parts of the final text. The author has reviewed and edited the output and takes full responsibility for the content of this publication. The author has read and agreed to the published version of the manuscript.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Dossa, K.F.; Miassi, Y.E. Global Policies and Perspectives of Biofuels and Bioenergy: Efforts to Combat Climate Change. In *Recent Trends in Lignocellulosic Biofuels and Bioenergy*; Springer: Cham, Switzerland, 2025; pp. 261–303.
2. Marks-Bielska, R.; Bielski, S.; Kurowska, K.; Zielińska-Chmielewska, A. First-Generation Biofuels vs. Energy Security: An Overview of Biodiesel and Bioethanol. *Energies* **2025**, *18*, 6055. [CrossRef]
3. Jeswani, H.K.; Chilvers, A.; Azapagic, A. Environmental sustainability of biofuels: A review. *Proc. R. Soc. A* **2020**, *476*, 20200351. [CrossRef] [PubMed]
4. Waters, K.; Altiparmak, S.O.; Shutters, S.T.; Thies, C. The Green Mirage: The EU's Complex Relationship with Palm Oil Biodiesel in the Context of Environmental Narratives and Global Trade Dynamics. *Energies* **2024**, *17*, 343. [CrossRef]
5. Nogales-Delgado, S. Biodiesel Production and Life Cycle Assessment: Status and Prospects. *Energies* **2025**, *18*, 3338. [CrossRef]
6. International Energy Agency (IEA). Biofuels Production—World Data and Projections. Available online: <https://www.iea.org/data-and-statistics/> (accessed on 12 May 2026).
7. U.S. Energy Information Administration (EIA). International Energy Statistics: Biodiesel Production. Available online: <https://www.eia.gov/international/data/world> (accessed on 12 May 2026).
8. OECD-FAO. OECD-FAO Agricultural Outlook 2023–2032 (Biofuels Chapter). Available online: https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agricultural-outlook_19991142 (accessed on 10 May 2026).
9. Eurostat. Production of Biofuels in the European Union. Available online: <https://ec.europa.eu/eurostat> (accessed on 11 May 2026).
10. REN21. Renewables 2023 Global Status Report. Available online: <https://www.ren21.net/reports/global-status-report/> (accessed on 9 May 2026).
11. Grand View Research. Biolubricants Market Size, Share and Trends Analysis Report. Available online: <https://www.grandviewresearch.com/industry-analysis/biolubricants-market> (accessed on 15 May 2026).
12. MarketsandMarkets. Biolubricants Market—Global Forecast to 2030. Available online: <https://www.marketsandmarkets.com/Market-Reports/biolubricants-market> (accessed on 15 May 2026).
13. Statista. Global Biolubricants Market Size from 2015 to 2030. Available online: <https://www.statista.com> (accessed on 14 March 2026).
14. European Commission. Bio-Based Products Market Data (Including Lubricants). Available online: <https://environment.ec.europa.eu/topics/circular-economy/> (accessed on 13 May 2026).
15. United States Department of Agriculture (USDA). BioPreferred Program—Biobased Lubricants. Available online: <https://www.biopreferred.gov> (accessed on 13 May 2026).
16. El Marji, L.; Sharara, M.; El Chakik, D.; Nakad, M.; Assaf, J.C.; Estephane, J. A Comparative Review of Biomass Conversion to Biodiesel with a Focus on Sunflower Oil: Production Pathways, Sustainability, and Challenges. *Processes* **2026**, *14*, 441. [CrossRef]
17. Nogales-Delgado, S.; Encinar, J.M.; González, J.F. A Review on Biolubricants Based on Vegetable Oils through Transesterification and the Role of Catalysts: Current Status and Future Trends. *Catalysts* **2023**, *13*, 1299. [CrossRef]
18. Demirbas, A. Biodiesel: A realistic fuel alternative for diesel engines. *Renew. Sustain. Energy Rev.* **2012**, *16*, 6441–6451. [CrossRef]
19. Knothe, G.; Krahel, J.; Van Gerpen, J. The biodiesel handbook. *Prog. Energy Combust. Sci.* **2017**, *58*, 36–59. [CrossRef]

20. Hoekman, S.K.; Broch, A.; Robbins, C.; Cenicerros, E.; Natarajan, M. Review of biodiesel composition, properties, and specifications. *Renew. Sustain. Energy Rev.* **2016**, *64*, 343–366. [\[CrossRef\]](#)
21. Bergwerf, J. MolView. Available online: <https://molview.org> (accessed on 14 May 2026).
22. Mueller, D.; Ferrão, M.F.; Marder, L.; Costa, A.B.; Schneider, R.C.S. Fourier Transform Infrared Spectroscopy (FTIR) and Multivariate Analysis for Identification of Different Vegetable Oils Used in Biodiesel Production. *Sensors* **2013**, *13*, 4258–4271. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Tehranizadeh, M.; Ghazanfari-Moghaddam, A.; Hashmipour Rafsanjani, H. A Weighted Principal Component Analysis Approach for Online Quality Determination of Biodiesel Using Spectrophotometry Data. *Energy Sources Part A* **2019**, *41*, 2214–2226. [\[CrossRef\]](#)
24. Santos, S.M.; Maciel, M.R.W.; Fregolente, L.V. Application of Multivariate Exploratory Techniques to Predict Kinematic Viscosity of Biodiesel from Vegetable and Algae Oils. *Chem. Eng. Trans.* **2022**, *92*, 739–744.
25. Jahirul, M.I.; Brown, R.J.; Senadeera, W. Correlation between Physicochemical Properties and Quality of Biodiesel. In *Application of Thermo-fluid Processes in Energy Systems*; Springer: Singapore, 2017; pp. 57–81.
26. Souza, V.M.G.; d’Avila, L.A.; da Cruz, I.O. HVO and biodiesel impact on diesel fuel stability: A multivariate data analysis approach. *Braz. J. Chem. Eng.* **2025**, *42*, 1113–1129. [\[CrossRef\]](#)
27. Boichenko, S.; Yakovlieva, A.; Zubenko, S.; Shkilniuk, I. Physical, Chemical, and Performance Properties of Biodiesel Fuels: A Comparative Study of Lipid-Based Feedstocks. *Energies* **2025**, *18*, 4274. [\[CrossRef\]](#)
28. Botella, L.; Bimbela, F.; Martín, L.; Arauzo, J.; Sánchez, J.L. Oxidation Stability of Biodiesel Fuels and Blends Using the Rancimat and PetroOXY Methods. Effect of 4-Allyl-2,6-Dimethoxyphenol and Catechol as Biodiesel Additives on Oxidation Stability. *Front. Chem.* **2014**, *2*, 43. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Canha, N.; Felizardo, P.; Correia, M.J.N. Controlling the Oxidative Stability of Biodiesel Using Oils or Biodiesel Blending or Antioxidants Addition. *Environ. Prog. Sustain. Energy* **2017**, *37*, 1031–1040. [\[CrossRef\]](#)
30. De Freitas, O.N.; Rial, R.C.; Cavalheiro, L.F.; Barbosa, J.M.D.S.; Nazário, C.E.D.; Viana, L.H. Evaluation of the Oxidative Stability and Cold Filter Plugging Point of Soybean Methyl Biodiesel/Bovine Tallow Methyl Biodiesel Blends. *Ind. Crops Prod.* **2019**, *140*, 111667. [\[CrossRef\]](#)
31. Hong, I.K.; Lee, J.R.; Lee, S.B. Fuel Properties of Canola Oil and Lard Biodiesel Blends: Higher Heating Value, Oxidative Stability, and Kinematic Viscosity. *J. Ind. Eng. Chem.* **2014**, *22*, 335–340. [\[CrossRef\]](#)
32. Karavalakis, G.; Stournas, S. Impact of Antioxidant Additives on the Oxidation Stability of Diesel/Biodiesel Blends. *Energy Fuels* **2010**, *24*, 3682–3686. [\[CrossRef\]](#)
33. Khan, E.; Ozaltin, K.; Spagnuolo, D.; Bernal-Ballen, A.; Piskunov, M.V.; Di Martino, A. Biodiesel from Rapeseed and Sunflower Oil: Effect of the Transesterification Conditions and Oxidation Stability. *Energies* **2023**, *16*, 657. [\[CrossRef\]](#)
34. Kpan, J.D.; Krah, J. The Impact of Adsorbents on the Oxidative Stability of Biodiesel and Its Influence on the Deterioration of Engine Oil. *Fuel* **2019**, *256*, 115984. [\[CrossRef\]](#)
35. Moser, B.R. Preparation of Fatty Acid Methyl Esters from Hazelnut, High-Oleic Peanut and Walnut Oils and Evaluation as Biodiesel. *Fuel* **2011**, *92*, 231–238. [\[CrossRef\]](#)
36. Ramos, M.J.; Fernández, C.M.; Casas, A.; Rodríguez, L.; Pérez, Á. Influence of Fatty Acid Composition of Raw Materials on Biodiesel Properties. *Bioresour. Technol.* **2008**, *100*, 261–268. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Schober, S.; Mittelbach, M. The Impact of Antioxidants on Biodiesel Oxidation Stability. *Eur. J. Lipid Sci. Technol.* **2004**, *106*, 382–389. [\[CrossRef\]](#)
38. Sui, M.; Li, F. Effect of TEPA on Oxidation Stability and Metal Ion Content of Biodiesel. *Renew. Energy* **2019**, *143*, 352–358. [\[CrossRef\]](#)
39. Suraj, C.K.; Krishnasamy, A.; Sundararajan, T. Investigations on Gradual and Accelerated Oxidative Stability of Karanja Biodiesel and Biodiesel–Diesel Blends. *Energy Fuels* **2019**, *33*, 9196–9204. [\[CrossRef\]](#)
40. Tang, H.; De Guzman, R.C.; Salley, S.O.; Ng, S.K. The Oxidative Stability of Biodiesel: Effects of FAME Composition and Antioxidant. *Lipid Technol.* **2008**, *20*, 249–252. [\[CrossRef\]](#)
41. Zhou, J.; Xiong, Y.; Liu, X. Evaluation of the Oxidation Stability of Biodiesel Stabilized with Antioxidants Using the Rancimat and PDSC Methods. *Fuel* **2016**, *188*, 61–68. [\[CrossRef\]](#)
42. Encinar, J.M.; Nogales, S.; González, J.F. Biodiesel and Biolubricant Production from Different Vegetable Oils through Transesterification. *Eng. Rep.* **2020**, *2*, e12190. [\[CrossRef\]](#)
43. Encinar, J.M.; Nogales-Delgado, S.; Pinilla, A. Biolubricant Production through Double Transesterification: Reactor Design for the Implementation of a Biorefinery Based on Rapeseed. *Processes* **2021**, *9*, 1224. [\[CrossRef\]](#)
44. Encinar, J.; Nogales-Delgado, S.; Sánchez, N.; González, J. Biolubricants from Rapeseed and Castor Oil Transesterification by Using Titanium Isopropoxide as a Catalyst: Production and Characterization. *Catalysts* **2020**, *10*, 366. [\[CrossRef\]](#)
45. Nogales-Delgado, S.; Encinar, J.M.; González, J.F. Safflower Biodiesel: Improvement of Its Oxidative Stability by Using BHA and TBHQ. *Energies* **2019**, *12*, 1940. [\[CrossRef\]](#)

46. Nogales-Delgado, S.; Guiberteau, A.; Encinar, J. Effect of tert-Butylhydroquinone on Biodiesel Properties during Extreme Oxidation Conditions. *Fuel* **2021**, *310*, 122339. [\[CrossRef\]](#)
47. Nogales-Delgado, S.; Martín, J.M.E. Cardoon Biolubricant through Double Transesterification: Assessment of Its Oxidative, Thermal and Storage Stability. *Mater. Lett.* **2021**, *302*, 130454. [\[CrossRef\]](#)
48. Nogales-Delgado, S.; Martín, J.M.E.; Ocaña, M.S. Use of Mild Reaction Conditions to Improve Quality Parameters and Sustainability during Biolubricant Production. *Biomass Bioenergy* **2022**, *161*, 106456. [\[CrossRef\]](#)
49. Nogales-Delgado, S.; Encinar, J.M.; González Cortés, Á. High Oleic Safflower Oil as a Feedstock for Stable Biodiesel and Biolubricant Production. *Ind. Crops Prod.* **2021**, *170*, 113701. [\[CrossRef\]](#)
50. UNE-EN 14214:2013 V2+A1:2018; Liquid Petroleum Products—Fatty Acid Methyl Esters (FAME) for Use in Diesel Engines and Heating Applications—Requirements and Test Methods. European Committee for Standardization: Brussels, Belgium, 2018.
51. Encinar, J.; Nogales-Delgado, S.; Álvarez-Medina, C. High Oleic Safflower Biolubricant through Double Transesterification with Methanol and Pentaerythritol: Production, Characterization, and Antioxidant Addition. *Arab. J. Chem.* **2022**, *15*, 103796. [\[CrossRef\]](#)
52. UNE-EN ISO 3104/AC:1999; Petroleum Products. Transparent and Opaque Liquids. Determination of Kinematic Viscosity and Calculation of Dynamic Viscosity. International Organization for Standardization: Geneva, Switzerland, 1999.
53. UNE-EN 14112; Fat and Oil Derivatives—Fatty Acid Methyl Esters (FAME)—Determination of Oxidation Stability (Accelerated Oxidation Test) 2017. European Committee for Standardization (CEN): Brussels, Belgium.
54. Jolliffe, I.T. *Principal Component Analysis*; Springer: New York, NY, USA, 2002.
55. Caldeira, C.; Freire, F.; Olivetti, E.A.; Kirchain, R. Fatty acid based prediction models for biodiesel properties incorporating compositional uncertainty. *Fuel* **2017**, *208*, 321–332.
56. Duong, S.; Lamharess-Chlaft, N.; Sicard, M.; Raepsaet, B.; Galvez, M.E.; Da Costa, P. New Approach for Understanding the Oxidation Stability of Neopolyol Ester Lubricants Using a Small-Scale Oxidation Test Method. *ACS Omega* **2018**, *3*, 10449–10459. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Filon, D.; Anastopoulos, G.; Karonis, D. Synthesis and Evaluation of Bio-Lubricants from Renewable Raw Materials. *Lubricants* **2024**, *12*, 446. [\[CrossRef\]](#)
58. Rahmadiawan, D.; Abrial, H.; Nasruddin, N.; Fuadi, Z. Stability, Viscosity, and Tribology Properties of Polyol Ester Oil-Based Biolubricant. *Int. J. Polym. Sci.* **2021**, *2021*, 5536047. [\[CrossRef\]](#)
59. Encinar, J.M.; Nogales, S.; González, J.F. The Effect of BHA on Oxidative Stability of Biodiesel from Different Sources. *Greenh. Gases Sci. Technol.* **2020**, *10*, 1193–1201. [\[CrossRef\]](#)
60. Nogales-Delgado, S.; Cabanillas, A.G.; Romero, Á.G.; Martín, J.M.E. Monitoring tert-Butylhydroquinone Content and Its Effect on a Biolubricant during Oxidation. *Molecules* **2022**, *27*, 8931. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Sanjurjo, C.; Rodríguez, E.; Viesca, J.L.; Battez, A.H. Influence of Molecular Structure on the Physicochemical and Tribological Properties of Biolubricants: A Review. *Lubricants* **2023**, *11*, 380. [\[CrossRef\]](#)
62. Schölkopf, B.; Smola, A.; Müller, K.R. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Comput.* **1998**, *10*, 1299–1319. [\[CrossRef\]](#)
63. McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* **2018**, arXiv:1802.03426. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.