

Modeling of Biodiesel Production via Transesterification using Inline Raman Spectroscopy^{*}

Ilias Bouchkira, Mohammad El Wajeh, Adel Mhamdi^{*}

*RWTH Aachen University, Process Systems Engineering (AVT.SVT),
Aachen 52074, Germany (*adel.mhamdi@avt.rwth-aachen.de).*

Abstract: We present a reaction kinetics model for biodiesel production via transesterification, which is calibrated using concentration measurements from inline Raman spectroscopy. The novel application of Raman spectroscopy in biodiesel production provides real-time monitoring of key reaction species, e.g., fatty acid methyl esters, triglycerides, methanol, and glycerol. We employ an automated semi-batch reactor to perform reaction experiments. A robust offline calibration process of the Raman device allows achieving high accuracy for concentration predictions ($R^2 = 0.99$). Moreover, using sodium methylate as the catalyst, we address a gap in the literature where kinetic parameter values for transesterification with this catalyst are unavailable. For accurate parameter estimation, we use genetic algorithms. A global sensitivity-based estimability analysis confirms the sufficiency of the experimental data. We determine confidence intervals through Hessian matrix estimation. Our model predictions are validated against experimental data at 60 °C, demonstrating excellent agreement. These results highlight the effectiveness of integrating Raman spectroscopy for modeling reaction kinetics, hence offering promising tools and models for monitoring, optimization, and control of biodiesel production processes.

Keywords: Biodiesel, Transesterification, Raman Spectroscopy, Kinetic Modeling, Parameter Estimation, Global Sensitivity Analysis.

1. INTRODUCTION

As the global demand for renewable and sustainable energy sources intensifies, biodiesel has emerged as a good alternative to fossil fuels. Its biodegradable and renewable properties, coupled with its potential to significantly reduce greenhouse gas emissions, make biodiesel an attractive solution in the quest for sustainable energy. Additionally, biodiesel compatibility with existing diesel engines without requiring substantial modifications underscores its viability as an alternative fuel. However, to fully exploit its potential, optimizing the biodiesel production process to enhance efficiency and cost-effectiveness is imperative (Prasad et al., 2024).

Biodiesel is commonly produced via the transesterification of triglycerides, a process that converts oils and fats into fatty acid methyl esters (FAME) and glycerol in the presence of an alcohol and a catalyst (El Wajeh et al., 2023). The efficiency of this process is heavily influenced by the choice of the oil, alcohol, and catalyst. While traditional catalysts such as sodium hydroxide (NaOH) and potassium hydroxide (KOH) are widely used, sodium methylate (NaOCH_3) has gained attention for its superior reactivity, higher conversion efficiency, and reduced risk of saponification, which can negatively influence biodiesel

production. The use of sodium methylate can lead to faster reaction times, and lower energy consumption.

To optimize biodiesel production, a deep understanding of the reaction kinetics is essential. Indeed, studies of reaction kinetics offer valuable insights into reaction mechanisms and rate-determining steps, which are critical for optimization of process design and operation. However, standard approaches for analysis of reaction kinetics such as Gas Chromatography (Tiyapongpattana et al., 2008), High-Performance Liquid Chromatography (Holčapek et al., 1999), Nuclear Magnetic Resonance (Ng and Yung, 2019) and others are often time-consuming and labor-intensive. In this context, Raman spectroscopy presents a promising approach to studying reaction kinetics, since it provides real-time, non-invasive access to concentrations of reacting species.

We investigate the use Raman spectroscopy for modeling the reaction kinetics of biodiesel production, particularly in sodium methylate-catalyzed transesterification. By integrating advanced spectroscopic techniques with modeling of reaction kinetics, this work seeks to enhance the accuracy and predictive capabilities of biodiesel production models. In this work, our primary focus lies in the development and validation of a reaction kinetics model specific to sodium methylate-catalyzed transesterification, since (to our knowledge) it is not available in the open literature. The application of Raman spectroscopy in this context represents a significant advancement in biodiesel

^{*} The authors gratefully acknowledge the financial support of the Kopernikus project SynErgie by the Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich.

production, providing a powerful tool for monitoring and optimizing reaction conditions, improving biodiesel yields, and reducing production costs.

The transesterification model involves six unknown parameters to be estimated using experimental data. Before proceeding to parameter estimation, we employ a global sensitivity-based estimability analysis to answer the question, whether the available measurements allow to identify all the parameters or only some of them.

The structure of this paper is as follows: Section 2 presents the experimental setup employed for biodiesel production. Section 3 outlines the methodologies used, including the transesterification modeling approach, Raman spectroscopy, and the procedures for parameter estimability and identification. Section 4 provides a detailed discussion of the key results, while Section 5 concludes the manuscript, summarizing the main findings.

2. EXPERIMENTAL SETUP

The experimental setup for biodiesel production through transesterification involves an arrangement of equipment and instruments as shown in Fig. 1. It is designed for an automated monitoring and control of the reaction process. The core of the setup is a 500 ml glass reactor (1), which is fitted with a jacket for both heating and cooling. This jacket is connected to a Julabo heating circulator (2), capable of achieving temperatures from $-20\text{ }^{\circ}\text{C}$ to $260\text{ }^{\circ}\text{C}$. Inside the reactor, a Pt100 temperature sensor (3) ensures accurate measurement of the reaction temperature.

To introduce the reactants into the reactor, two LabDos pumps (4)–(5) are utilized: one for the oil and another for methanol and catalyst mixture. Each reactant is stored in a separate tank (6)–(7). The reactor is equipped with a Heidolph stirrer (8) to ensure thorough mixing of the reactants, which is crucial for reaction efficiency.

A Raman spectroscopy probe (9) is put inside the reactor to capture the specific vibrational signatures of the species involved in the reaction and thus enables inline and real-time monitoring of the reaction progress. This probe sends the measured spectra to the Raman Spectrometer (11) operating within the $100\text{--}3200\text{ cm}^{-1}$ wavelength range, allowing for continuous assessment of the chemical composition of the mixture. The entire system is automated and controlled via the LabManager system by Hitec Zang (10), which interfaces with a custom-developed process interface in the LabVision software (12). This integration provides direct data acquisition and monitoring, ensuring precise tracking, and analysis of the reaction.

3. METHODS

3.1 Transesterification Modeling

We investigate the transesterification of refined oil using methanol in the presence of sodium methylate as a catalyst in a semi-batch reactor operation. The primary objective is to model the kinetics of the transesterification reaction, which involves the conversion of triglycerides (TG) by methanol (MeOH) into diglycerides (DG), monoglycerides (MG), FAME, and glycerol (G) through intermediate

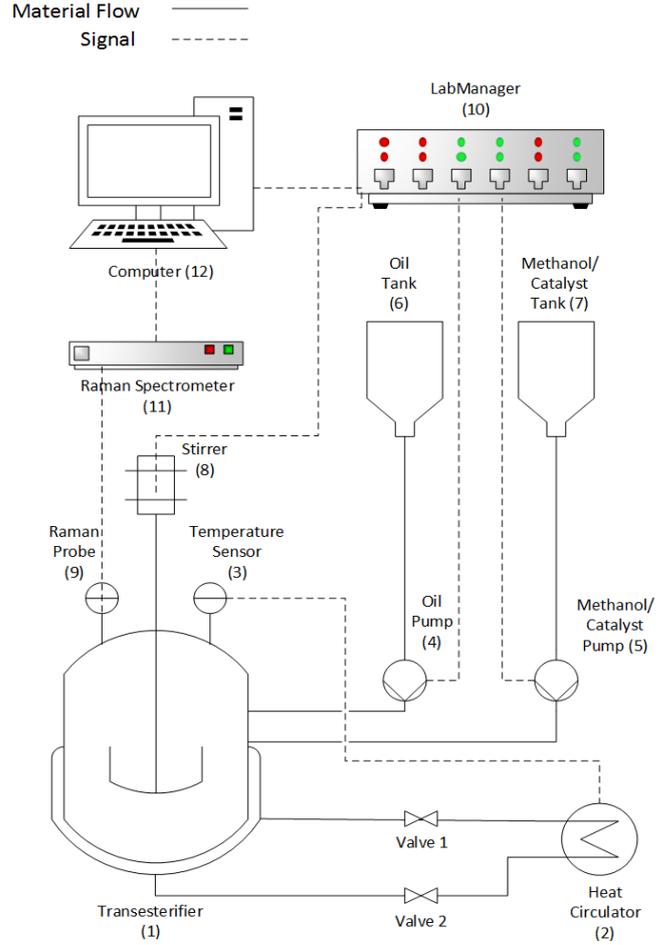
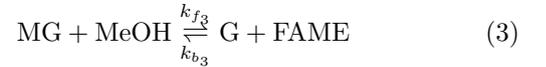
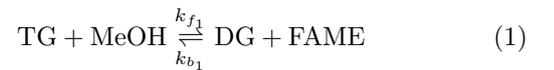


Fig. 1. Experimental setup used for oil transesterification to measure concentrations of reaction species.

steps. In the model, we assume the following three-step reaction mechanism (Azad et al., 2024; El Wajeh et al., 2023):



where, k_{f_j} and k_{b_j} represent the forward and reverse reaction rate constants for each step $j \in \{1, 2, 3\}$, respectively.

The net reaction rates r_j are given below,

$$r_1 = k_{f_1} C_{\text{TG}} C_{\text{MeOH}} - k_{b_1} C_{\text{FAME}} C_{\text{DG}} \quad (4)$$

$$r_2 = k_{f_2} C_{\text{DG}} C_{\text{MeOH}} - k_{b_2} C_{\text{FAME}} C_{\text{MG}} \quad (5)$$

$$r_3 = k_{f_3} C_{\text{MG}} C_{\text{MeOH}} - k_{b_3} C_{\text{FAME}} C_{\text{G}} \quad (6)$$

where C_i represents the concentration of species i in the reactor.

Initially (at $t = 0$ s), the reactor contains a volume V_0 corresponding to 300 g of refined oil (assumed to be 100%) triglyceride. The content is heated to 60 °C, and a stoichiometric mixture of methanol and catalyst is then pumped into the reactor with a mass flow rate of 2.5 g/min during about 15 minutes. We assume that the thermophysical properties of the mixture do not change over time.

The material balances for each species are:

$$\frac{dN}{dt} = \dot{N}_{in} \quad (7)$$

$$\frac{dn_{TG}}{dt} = -r_1 V \quad (8)$$

$$\frac{dn_{DG}}{dt} = (r_1 - r_2) V \quad (9)$$

$$\frac{dn_{MG}}{dt} = (r_2 - r_3) V \quad (10)$$

$$\frac{dn_{MeOH}}{dt} = \dot{n}_{in, MeOH} - (r_1 + r_2 + r_3) V \quad (11)$$

$$\frac{dn_{FAME}}{dt} = (r_1 + r_2 + r_3) V \quad (12)$$

$$\frac{dn_G}{dt} = r_3 V \quad (13)$$

$$\frac{dn_{Catalyst}}{dt} = \dot{n}_{in, Catalyst} \quad (14)$$

$$(15)$$

The reactor volume V is computed as follows:

$$V = \sum_i \frac{n_i \cdot MW_i}{\rho_i} \quad (16)$$

where n_i is the molar amount, MW_i is the molar mass, and ρ_i is the mass density of each species in the reactor, i.e., TG, DG, MG, MeOH, FAME, glycerol, and NaOCH₃. Note that all the involved thermodynamic properties shown in Table 1 are available in the NIST database (Linstrom and Mallard, 2024).

Table 1. Properties of the transesterification reaction reagents and products.

Species	Molar mass (g/mol)	Density (g/l)
TG	885	900
DG	620	920
MG	355	950
MeOH	32	791
FAME	292	880
Glycerol	92	1260
NaOCH ₃	54.03	945

3.2 Raman Spectroscopy

To ensure the accuracy of the Raman calibration model that computes concentrations from the measured spectra, we conduct an offline calibration process. Reference samples with known concentrations of FAME, oil, methanol, glycerol, and sodium methylate are prepared, and their Raman spectra are measured. The spectral data undergo

preprocessing, including baseline correction, normalization, and smoothing, to enhance signal quality and minimize noise. We process the measured spectra using the PEAXACT software (SPACT-GmbH, 2023), applying a Partial Least Squares (PLS) regression model to establish the correlation between spectral intensity and mass fractions. A calibration curve is then constructed based on this regression analysis, providing a quantitative relationship between the Raman signal and species concentrations.

To ensure broad applicability and robustness, the calibration covers a wide range of concentrations by preparing pure, binary, ternary, and quaternary mixtures while carefully selecting compositions that prevent unintended reactions. The experimental design follows a systematic approach where calibration points are evenly distributed across relevant concentration ranges to improve predictive accuracy. The model’s performance is evaluated through cross-validation and residual analysis, ensuring high predictive capability.

This structured approach to sample preparation, preprocessing, and calibration ensures that the Raman model reliably monitors the composition of the reaction mixture, thereby improving the accuracy of concentration predictions and enhancing the interpretation of the results presented in Fig. 2.

3.3 Estimability Analysis

We address the estimability of the model unknown parameters using our global-sensitivity based estimability analysis approach (Bouchkira et al., 2021, 2024). It is based on computing the Fisher information matrix using global sensitivities of the model unknown parameters with respect to the measured outputs. The latter is used to calculate the estimability magnitudes of the unknown parameters, which can be compared to estimability criteria and conclude if the used experimental data contains enough information to identify the parameters accurately. The approach has been successfully employed in several previous works (Bouchkira et al., 2023, 2022)

3.4 Parameter Estimation

The unknown values of the model parameters are estimated from the concentration measurements by minimizing the following objective function:

$$\mathcal{L} = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{C_{i,j}^{exp} - C_{i,j}^{mod}}{C_{i,j}^{exp}} \right)^2 \quad (17)$$

where $C_{i,j}^{exp}$ is the experimental concentration of species i at time j and $C_{i,j}^{mod}$ is the modeled concentration of species i at time j . n is the number of species in the system and m is the number of sampling times.

We implement the optimization problem in MATLAB (MathWorks, 2023) and solve it using its local and global solvers, such as the genetic algorithm (GA), as it handles complex, non-convex optimization problems efficiently. Additionally, parallel computing is employed to expedite the process by distributing the workload, significantly reducing computation time.

The confidence intervals for the estimated parameters were computed using the Hessian \mathbf{H} , with its H_{ij} element given by:

$$H_{ij} = \frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_j} \quad (18)$$

Where θ_i and θ_j are the parameters being estimated. The covariance matrix is then approximated as the inverse of the Hessian matrix:

$$\mathbf{Cov}(\hat{\theta}) = \mathbf{H}^{-1} \quad (19)$$

The diagonal elements of the covariance matrix correspond to the variances of the estimated parameters. The standard error for each parameter is computed as the square root of the corresponding diagonal element:

$$SE(\hat{\theta}_i) = \sqrt{\mathbf{Cov}_{ii}(\hat{\theta})} \quad (20)$$

where $SE(\hat{\theta}_i)$ is the standard error of the i th parameter, and $\mathbf{Cov}_{ii}(\hat{\theta})$ is the i th diagonal element of the covariance matrix. To construct the 95 % confidence intervals for each parameter, we use the following equation:

$$\hat{\theta}_i \pm \zeta_{st} \cdot SE(\hat{\theta}_i) \quad (21)$$

ζ_{st} corresponds to the critical value of the standard normal distribution for a 95 % confidence level. It is taken from the student table based on the number of unknown parameters and the number of freedom degrees.

4. RESULTS AND DISCUSSION

4.1 Raman Spectrum Calibration

Fig. 2 presents the calibration results, highlighting the comparison between the measured compositions of the prepared samples and the predictions from the Raman calibration model. The calculated R^2 for the calibration is 0.99, indicating a very good prediction accuracy. These high R^2 values confirm that the Raman calibration model is highly reliable in predicting the composition of the reaction mixture across the entire range of tested concentrations. The calibration results indicate that the Raman calibration model is well-suited for real-time monitoring, providing precise and consistent measurements for an effective determination of the reaction kinetics.

4.2 Estimability Analysis

Fig. 3 shows the ranking of the unknown parameters according to their estimability using the global estimability analysis toolbox (Bouchkira et al., 2024). The estimability magnitudes are compared to an estimability criterion ($\zeta = 0.04$) by Zhang et al. (2015), which indicates that for a parameter to be estimable from an available output, a change of 2% of its nominal value should cause at least 10% of a change on the model output. In our case, it can be seen that the least estimable parameter has an estimability magnitude of 0.127, meaning that the available database of experiments is sufficient to accurately estimate all model parameters.

4.3 Parameter Estimation Results

Table 2 summarizes the identification results. It shows the optimal values of the estimated parameters along with their confidence intervals.

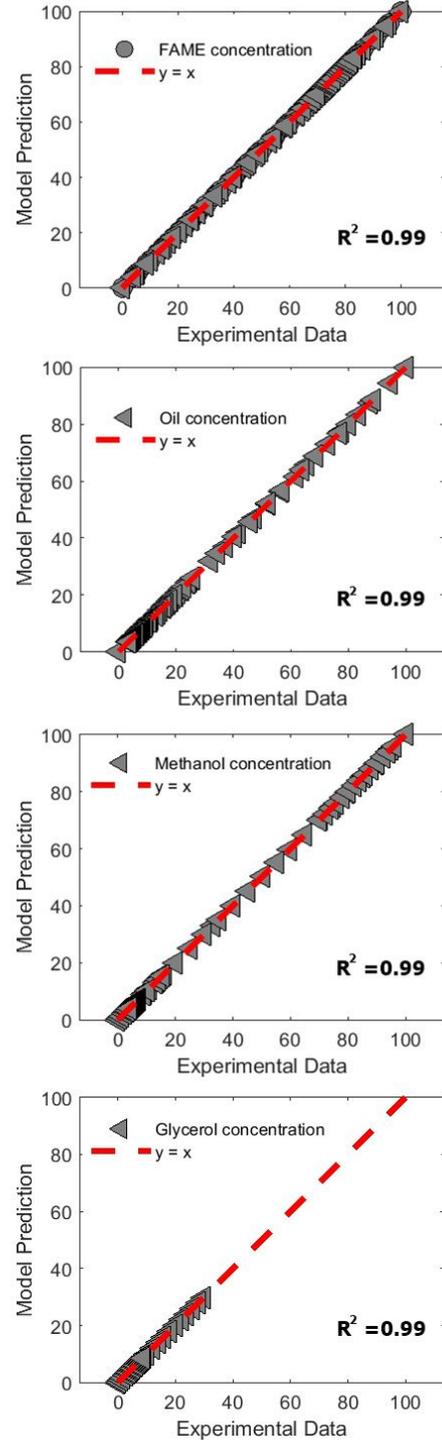


Fig. 2. Raman spectrum-based model calibration results (predictions vs experiments).

Fig. 4 and Table 3 show the concentration evolution over time during the transesterification reaction at 60 °C. We observe a good overall alignment between experimental, simulated and validation data (not used in the identification process) for all measured species. The production of FAME follows a steady rise, reaching a plateau near 4000 seconds, indicating the model's success in predicting the kinetics of biodiesel formation. The formation of glycerol is also well predicted, with both experimental and simulated data indicating smooth accumulation. Similarly,

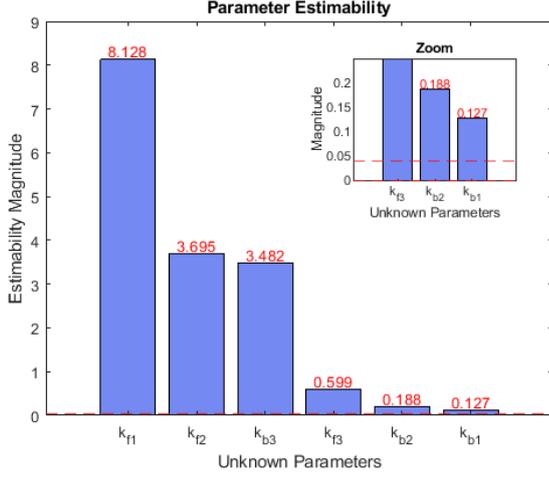


Fig. 3. Estimability analysis results (ranking of the unknown parameters according to their estimability magnitudes).

Table 2. Values of parameters and their 95% confidence intervals.

Parameter $\hat{\theta}_i$	Value at 60°C	Confidence Interval
k_{f1}	1.1452	[1.1451, 1.1453]
k_{b1}	25.711	[25.709, 25.713]
k_{f2}	0.1142	[0.1141, 0.1143]
k_{b2}	0.0137	[0.0136, 0.0138]
k_{f3}	31.586	[31.585, 31.587]
k_{b3}	0.0140	[0.0139, 0.0141]

the decline in triglyceride (oil) concentration reflects efficient conversion during the reaction, closely matching experimental observations. The consumption of methanol shows a distinct pattern in which its concentration initially increases as it is added to the system, followed by a gradual decline as it is consumed in the reaction.

The consistency across species in the simulated model at 60°C suggests that the kinetic parameters used are well-calibrated for this temperature.

We observe a slight discrepancy between the experimental and simulated profiles for methanol. This may be due to experimental variations or limitations in capturing the precise mass transfer rates within the reactor system. This issue will be investigated in a future work.

The unmeasured species diglycerides and monoglycerides form and deplete during the transesterification reaction. Simultaneously, the catalyst concentration increases during the injection of the methanol-catalyst mixture and stabilizes once the stoichiometric amount has been fully introduced. The total volume of the mixture inside the transesterifier follows an ideal behavior, as no volume excess was considered for the sake of simplification.

The kinetic parameters for the three stages of the transesterification of triolein into FAME and glycerol reveal important insights into the reaction dynamics. In the first stage, the forward rate (k_{f1}) of converting triglycerides

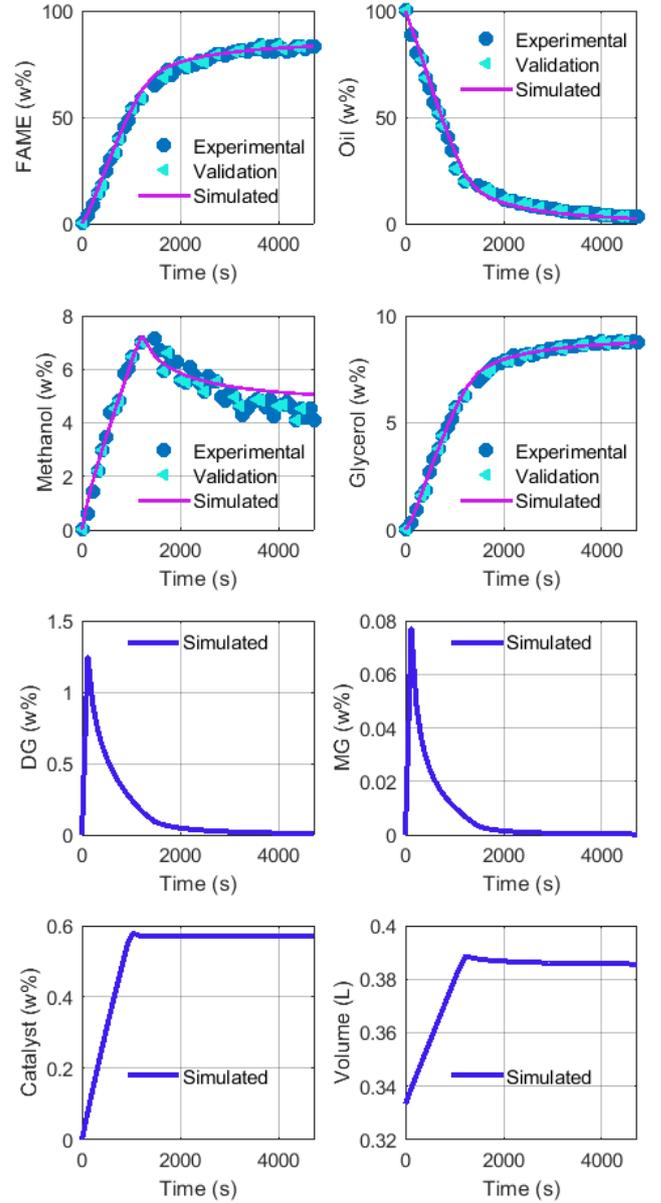


Fig. 4. Prediction of the model against experimental measurements at T=60°C.

into diglycerides is relatively low compared to the backward rate (k_{b1}), indicating that the reaction favors the reverse direction. This suggests that achieving a complete conversion of triglycerides requires optimizing reaction conditions, such as using an excess of methanol or ensuring efficient mixing to drive the reaction forward, as is usually done in industry. The slower forward reaction, combined with a high reverse rate, implies that the equilibrium at this stage tends to favor the triglyceride form, making it a critical point to focus on for improving overall yield.

The second and third stages show contrasting dynamics. In the second stage, the forward reaction (k_{f2}) is slow but significantly favored over the backward reaction (k_{b2}), indicating that diglycerides convert to monoglycerides more readily, though at a slower rate. This step may represent a bottleneck in the process. In the third stage, the forward

rate (k_{f_3}) is extremely fast compared to the backward rate (k_{b_3}), suggesting that once monoglycerides are formed, they quickly and almost irreversibly convert into glycerol and FAME. This rapid, forward-favored reaction helps ensure the efficient production of glycerol and FAME, which is advantageous in industrial settings where maximizing product formation is crucial.

Table 3. Species mass fraction in time. 1: FAME, 2: Oil, 3: Methanol, 4: Glycerol, 5: Diglycerides, 6: Monoglycerides, 7: Sodium methylate.

Time (s)	1	2	3	4	5	6	7
0	0	100	0	0	0	0	0
320	14.32	81.18	2.63	1.49	0.16	0.01	0.20
580	29.38	62.94	4.14	3.08	0.10	0.01	0.36
860	44.75	44.33	5.64	4.70	0.06	0.00	0.51
1220	61.24	24.28	7.43	6.43	0.03	0.00	0.58
1660	72.19	13.38	6.26	7.58	0.01	0.00	0.58
1940	75.49	10.09	5.90	7.93	0.01	0.00	0.58
2200	77.58	8.00	5.68	8.15	0.01	0.00	0.58
2480	79.22	6.38	5.50	8.32	0.01	0.00	0.58
2740	80.37	5.23	5.38	8.44	0.00	0.00	0.58
3020	81.32	4.28	5.28	8.54	0.00	0.00	0.58
3280	82.03	3.58	5.20	8.61	0.00	0.00	0.58
3560	82.64	2.97	5.13	8.68	0.00	0.00	0.58
3820	83.10	2.50	5.08	8.73	0.00	0.00	0.58
4100	83.51	2.10	5.04	8.77	0.00	0.00	0.58
4360	83.83	1.78	5.01	8.80	0.00	0.00	0.58
4640	84.11	1.50	4.98	8.83	0.00	0.00	0.58

5. CONCLUSION

This study develops and validates a kinetic model for biodiesel production via transesterification, providing in-line Raman spectroscopy for precise monitoring of reaction species. The use of sodium methylate as a catalyst marks a key advancement, as the kinetic parameters for this catalyst are previously unavailable in the literature. Through robust offline calibration and advanced parameter estimation using genetic algorithms, the model demonstrates good predictive accuracy for species concentrations. The integration of Raman spectroscopy enables accurate real-time data acquisition, making it a valuable tool for real-time process control in biodiesel production.

Future work will focus on refining the model by incorporating additional physical phenomena such as methanol evaporation at higher temperatures, which is identified as a source of deviation in model predictions of methanol. Additionally, applying this model to other catalysts and reaction conditions provides broader insights into biodiesel production kinetics. Further exploration of Raman spectroscopy in other complex chemical systems and coupling this kinetic model with model-based approaches holds significant potential for improving industrial biodiesel production efficiency and sustainability.

6. ACKNOWLEDGMENT

The authors express their gratitude to Martin Kleber and Wolfgang Hoffmann from Cargill GmbH for supplying the chemicals and for the valuable discussions that significantly enhanced the quality of this work.

REFERENCES

- Azad, A.K., Jadeja, A.C., Doppalapudi, A.T., Hassan, N.M.S., Nabi, M.N., and Rauniyar, R. (2024). Design and simulation of the biodiesel process plant for sustainable fuel production. *Sustainability*, 16(8), 3291.
- Bouchkira, I., Benjelloun, S., Khamar, L., and Latifi, A.M. (2023). Thermodynamic modeling and parameter estimability analysis of a wet phosphoric acid process with impurities. *Fluid Phase Equilibria*, 564, 113594.
- Bouchkira, I., Latifi, A.M., and Benyahia, B. (2024). Estant—a toolbox for standardized and effective global sensitivity-based estimability analysis. *Computers & Chemical Engineering*, 186, 108690.
- Bouchkira, I., Latifi, A.M., Khamar, L., and Benjelloun, S. (2021). Global sensitivity based estimability analysis for the parameter identification of pitzer’s thermodynamic model. *Reliability Engineering & System Safety*, 207, 107263.
- Bouchkira, I., Latifi, A.M., Khamar, L., and Benjelloun, S. (2022). Modeling and multi-objective optimization of the digestion tank of an industrial process for manufacturing phosphoric acid by wet process. *Computers & Chemical Engineering*, 156, 107536.
- El Wajeh, M., Mhamdi, A., and Mitsos, A. (2023). Dynamic modeling and plantwide control of a production process for biodiesel and glycerol. *Industrial & Engineering Chemistry Research*, 62(27), 10559–10576.
- Holčapek, M., Jandera, P., Fischer, J., and Prokeš, B. (1999). Analytical monitoring of the production of biodiesel by high-performance liquid chromatography with various detection methods. *Journal of chromatography A*, 858(1), 13–31.
- Linstrom, P. and Mallard, W. (2024). Nist chemistry webbook, nist standard reference database number 69. <https://webbook.nist.gov/chemistry/>. National Institute of Standards and Technology, Gaithersburg MD, 20899.
- MathWorks, I. (2023). *MATLAB: The Language of Technical Computing*. MathWorks, Natick, Massachusetts. Version R2023a.
- Ng, M.H. and Yung, C.L. (2019). Nuclear magnetic resonance spectroscopic characterisation of palm biodiesel and its blends. *Fuel*, 257, 116008.
- Prasad, S., Yadav, K.K., Kumar, S., Pandita, P., Bhutto, J.K., Alreshidi, M.A., Ravindran, B., Yaseen, Z.M., Osman, S.M., and Cabral-Pinto, M.M. (2024). Review on biofuel production: Sustainable development scenario, environment, and climate change perspectives- a sustainable approach. *Journal of Environmental Chemical Engineering*, 12(2), 111996.
- SPACT-GmbH (2023). Peaxact software for quantitative spectroscopy. <https://www.s-pact.de/>. Last accessed on 2024-10-22.
- Tiyapongpattana, W., Wilairat, P., and Marriott, P.J. (2008). Characterization of biodiesel and biodiesel blends using comprehensive two-dimensional gas chromatography. *Journal of separation science*, 31(14), 2640–2649.
- Zhang, X.Y., Trame, M.N., Lesko, L.J., and Schmidt, S. (2015). Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models. *CPT: pharmacometrics & systems pharmacology*, 4(2), 69–79.