

Application of Uncertainty and Sensitivity Analysis to a Kinetic Model for Enzymatic Biodiesel Production

Jason Price*, Mathias Nordblad*, John M. Woodley*, Jakob K. Huusom**

*Center for Process Engineering and Technology, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark (e-mail: {japr,matn,jw}@kt.dtu.dk).

**CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søtofts Plads, Building 229, DK-2800 Kgs. Lyngby, Denmark (e-mail: jkh@kt.dtu.dk)

Abstract: This paper demonstrates the added benefits of using uncertainty and sensitivity analysis in the kinetics of enzymatic biodiesel production. For this study, a kinetic model by Fedosov and co-workers is used. For the uncertainty analysis the Monte Carlo procedure was used to statistically quantify the variability in the model outputs due to uncertainties in the parameter estimates; showing the model is most reliable in the start (first 5 hours) of the reaction. To understand which input parameters are responsible for the output uncertainty, two global sensitivity methods (Standardized Regression Coefficients, and Morris screening) were used. The results from both sensitivity analyses identified that only 10 of the 32 parameters are influential to the model outputs. The model was then simplified by removing the non-influential parameters. A parity plot of the simplified model vs. the full model gave a R^2 value of over 0.95 for all the model outputs.

Keywords: Modelling, Sensitivity Analysis, Monte-Carlo Simulations, Enzymatic Biodiesel

1. INTRODUCTION

In order to determine the optimal conditions to operate the enzymatic biodiesel process, numerous experiments are done to help gain a better understanding of the system. While experiments are necessary to gain understanding, they are costly in terms of time and resource investments. Our viewpoint is that modelling can be a valuable tool to help focus the experimental work need for process understanding and to support further process development. A model of the system can be used to simulate potential experiments and also used to optimize the process. Integral to the simulation and optimisation step is the availability of reliable models.

Descriptions of the various kinetic models for enzymatic transesterification of vegetable oils are quite numerous (Al-Zuhair (2005), Pilarek and Szweczyk (2007), Cheirsilp *et al.* (2008), Calabrò *et al.* (2010), Li *et al.* (2010), Lv *et al.* (2010), Fedosov *et al.* (2012)). Depending on the process modelling goals, one model may offer a particular advantage over another. In terms of process design, it is desired that the kinetics describing enzymatic biodiesel production can predict the concentration of the major species in the reaction and how the process responds to changes in:

- Alcohol/oil molar ratio
- Water and FFA concentrations
- Different enzyme loading and stability
- Temperature
- Mixing (interfacial area of the oil–water interface)

One shortfall with the existing kinetic models is that, to the best of our knowledge, none that describe the enzymatic transesterification have been statistically analysed to ascertain the working bounds of the model. This remains as a

weak point in the credibility of these kinetic models and hence their applicability for engineering design purposes such as reactor configuration, optimisation and control. The aforementioned points set the stage for this study.

When modelling a process, the model parameters (e.g. kinetic parameters) and process conditions (e.g. flow rate or concentration) are the main components that influence the prediction quality of a model (Campiongo and Saltelli (1997)). In a host of various engineering fields, the standard Monte Carlo procedure has been used to statistically analyse the effect of uncertainty in the input factors (model parameters and/or process conditions) on the model outputs (uncertainty analysis); along with sensitivity analysis based on variance decomposition to identify and quantify which input factors were most influential to the model outputs (Sin *et al.* (2009), Vangsgaard *et al.* (2012), Saltelli *et al.* (2006)).

In this paper the kinetic model presented by Fedosov *et al.* (2012) is used as a case study. The aim is to:

- Evaluate the working bounds of a chosen model, for a predetermined parameter uncertainty (Uncertainty Analysis – Monte Carlo simulations). The outcome is a better understanding of the predictive accuracy of the model during the course of reaction. This gives insight into how well the model predicts the concentrations of key compounds; which forms the basis for any model based decision making for process design and operation
- Identify the group of parameters that are insensitive to the model outputs (Sensitivity Analysis - Standard Regression Coefficients and Morris screening). The outcome is a better understanding of which subsets of parameters are influential and non-influential. The influential parameters

help identify which parameters contribute most to the variance in the predicted concentrations of the model outputs. From which it can be deduced what mechanisms dominates at a particular point in the reaction. The non-influential parameters have negligible contribution to the variance in the model outputs, and can be fixed within the range of the parameter variability, aiding in model simplification.

The article is organised as follows. The methodology used is introduced. The case study is then presented along with the methods used in the simulations. The results from the uncertainty analysis are discussed, followed by the discussion of the results from the sensitivity analysis of the Standard Regression Coefficients and Morris screening. The results of the uncertainty and sensitivity analysis are then put into perspective for use in further engineering work

2. METHODOLOGY

To help improve the modelling process Foss *et al.* (1998) investigated the process of model development in chemical industries and laid out some guidelines to improve modelling technology. More recently Heitzig *et al.* (2011) proposed a generic methodology that structures the process of model development and analysis. This coupled with the work done by Sin *et al.* (2009) where statistical tools are used during the modelling process are combined in the methodology used in this work. The proposed methodology is illustrated in Fig. 1.

Defining the modelling objectives is essential in framing the goals and expected outcomes from the model. The information gathering process entails the collection of relevant experimental data along with phenomena occurring in the system such as reaction kinetics. To aid in this step, advice from experts, experimental observations and literature reviews are essential. During the model construction phase, modelling assumptions were made to help frame the problem. Mass and energy balances are made around the system boundaries and the constitutive equations relating to the

underling process phenomena are mathematically formulated. For the uncertainty analysis, the standard Monte Carlo procedure is used to propagate and analyse the uncertainty in the model parameters. To evaluate and rank the output variance of the model with respect to the model parameters, Standard Regression Coefficients and Morris screening is used for the sensitivity analysis. Two sensitivity analysis methods are used, to make it possible to corroborate the results obtained. If both methods identify the same parameters then it gives greater confidence in the results obtained.

3. CASE STUDY: KINETIC MODELLING OF ENZYMATIC BIODIESEL PRODUCTION

As mentioned in the introduction various kinetic models for enzymatic transesterification of vegetable oils have been proposed (Al-Zuhair (2005), Pilarek and Szewczyk (2007), Cheirsilp *et al.* (2008), Calabrò *et al.* (2010), Li *et al.* (2010), Lv *et al.* (2010), Fedosov *et al.* (2012)). The kinetic model by Fedosov *et al.* (2012) however, is interesting to work with, given the characteristics of the kinetic model (Table 1), provides specific outcomes that are necessary to aid in process development and is hence used in this study.

3.1. Kinetic model overview

Fedosov *et al.* (2012) presents a model for Novozym 435 (immobilized *Candida Antarctica* lipase B) for the conversion of rapeseed oil to biodiesel using methanol. In the model it is assumed reversibility of all reaction steps, and the phase boundaries in the system are imitated by introducing various molecular species of water and methanol. The model also contains competitive alcohol inhibition. The kinetic parameters are taken from the authors work and further description of the model can be found in Fedosov *et al.* (2012). For completeness the reaction scheme along with the kinetic parameters are presented in Table 2.

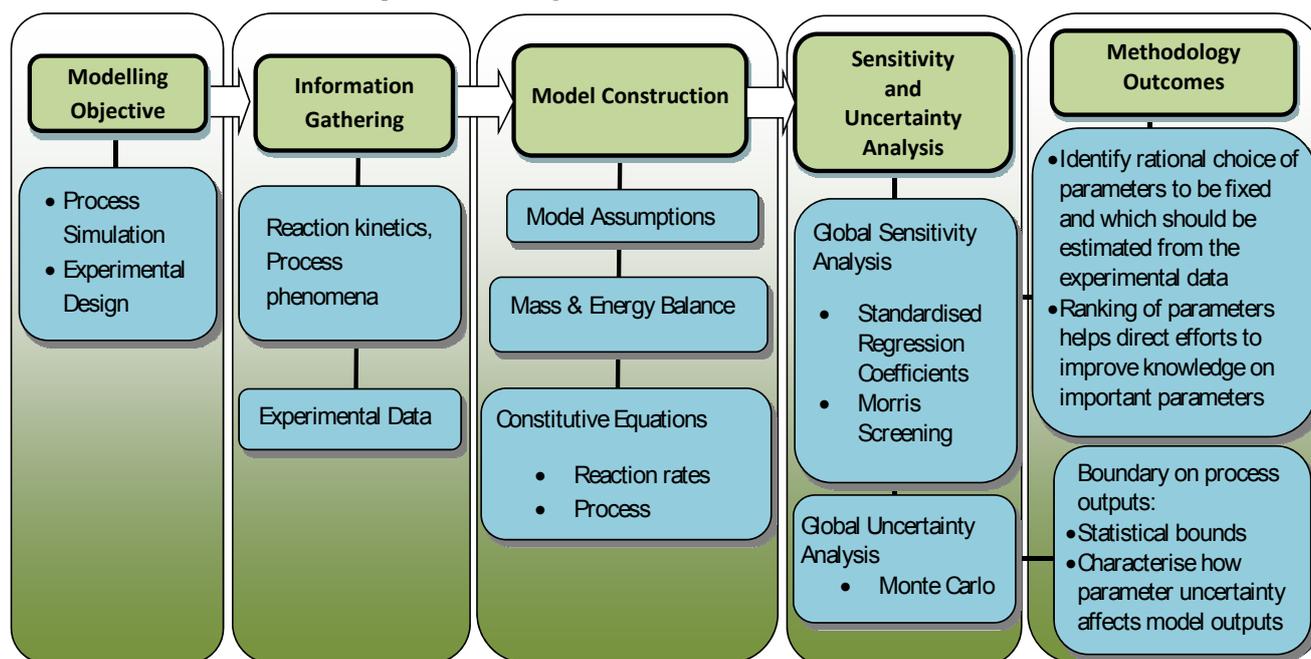


Fig. 1. Methodology used in this work for the Sensitivity and Uncertainty analysis.

Table 1 Potential uses of the kinetic model by Fedosov et al. (2012) given the model characteristics

Kinetic Model Characteristics	Modelling Outcomes
Kinetics describing how the reactants and products of interest (TAG, DAG, MAG, FFA, FFAE, Water, Glycerol, and Alcohol) varies during the entire reaction.	For a given alcohol/oil molar ratio determine when the reaction is complete
	Estimate how the changes in water and FFA concentrations affect the course of the reaction
Kinetic model includes the enzyme concentration in its mathematical expression	Estimate the required enzyme concentration to achieve a desired biodiesel yield in a specific time

The main reactions for the kinetic scheme presented in Table 2 are, esterification and transesterification. The free fatty acids (FFA, F) are esterified to fatty acid alkyl esters (FAAE, B) and in the reverse reaction the alkyl esters (biodiesel) are hydrolysed to FFA. The transesterification of the acylglycerides with methanol (CH, CN is the "aggregated" form of CH) is modelled as reversible steps, where, triglyceride (TAG, T) is converted to diglyceride (DAG, D), DAG is converted to monoglyceride (MAG, M) and MAG is converted to glycerol (G). Each conversion step yields one FFAE molecule, giving a total of three FFAE per TAG molecule.

In Table 2 the first 6 reaction represents the cyclic conversions between E (free enzyme) and EX (enzyme with the conjugated fatty acid). To imitate decreasing hydrolytic activity at increasing concentration of water, various species for water are introduced (W as water as a fine micella, WW and WN). $W > WW > WN$ at increasing concentrations of water. The bindings and equilibriums, represent the equilibrium between the enzyme and a substrate or the

Table 2 Kinetic reactions along with the parameters used in the simulation (Fedosov et al. (2012))

i	Reactions	Rate of reaction (r_i)	Parameters (value)		Differential Eqns.
			($M^{-1} min^{-1}$)	($M^{-1} min^{-1}$)	
1	$T \leftrightarrow \dots D$	$k_1 [T][E] - k_2 [EX][D]$	(k1)14	(k2)2.0	$\frac{d[T]}{dt} = -(r_1 + r_{20})$
2	$D \leftrightarrow \dots M$	$k_3 [D][E] - k_4 [EX][M]$	(k3)700	(k4)125	$\frac{d[D]}{dt} = r_1 - r_2$
3	$M \leftrightarrow \dots G$	$k_5 [M][E] - k_6 [EX][G]$	(k5)725	(k6)153	$\frac{d[M]}{dt} = r_2 - r_3$
4	$B \leftrightarrow \dots CH$	$k_7 [B][E] - k_8 [EX][CH]$	(k7)180	(k8)2000	$\frac{d[B]}{dt} = -r_4$
5	$F \leftrightarrow \dots W$	$k_9 [F][E] - k_{10} [EX][W]$	(k9)1500	(k10)2600	$\frac{d[F]}{dt} = -(r_5 + r_6)$
6	$F + W \leftrightarrow \dots WW$	$k_{11} [F][E][W] - k_{12} [EX][WW]$	(k11)10	(k12)490	$\frac{d[G]}{dt} = r_3 - r_{13}$
	Bindings and equilibriums		($M^{-n} min^{-1}$)	(min^{-1})	$\frac{d[W]}{dt} = r_5 - r_6 - 2r_7$
7	$2 \cdot W \leftrightarrow WW$	$k_{17} [W][W] - k_{18} [WW]$	(k17)100	(k18)17	$\frac{d[WW]}{dt} = r_6 + r_7 + r_9 + r_{10} - r_8$
8	$WW + T \rightarrow WN + T$	$k_{19} [WW][T]$	(k19)500	-	$\frac{d[WN]}{dt} = r_8 + r_{19} - r_9 - r_{10} - r_{17} - r_{18}$
9	$WW \leftarrow WN$	$k_{20} [WN]$	-	(k20)15	$\frac{d[CH]}{dt} = 2r_{12} + r_4 - 2r_{11} - r_{14}$
10	$WW + F \leftarrow WN + F$	$k_{24} [WN][F]$	-	(k24)30	$\frac{d[CN]}{dt} = r_{11} - r_{12} - 2r_{15} - r_{16}$
11	$2 \cdot CH + T \rightarrow CN + T$	$k_{26} [CH][CH][T]$	(k26)100	-	$\frac{d[E]}{dt} = r_{19} - r_1 - r_2 - r_3 - r_4 - r_5 - r_6 - r_{13} - r_{14} - r_{15} - r_{16} - r_{17} - r_{18} - r_{20}$
12	$2 \cdot CH \leftarrow CN$	$k_{27} [CN]$	-	(k27)82	$\frac{d[EX]}{dt} = -r_1 + r_2 + r_3 + r_4 + r_5 + r_6$
13	$E + G \leftrightarrow EG$	$k_{15} [G][E] - k_{16} [EG]$	(k15)10	(k16)6.2	$\frac{d[ET]^*}{dt} = r_{20}$
14	$E + CH \leftrightarrow EC$	$k_{28} [CH][E] - k_{29} [ECH]$	(k28)100	(k29)140	$\frac{d[EG]^*}{dt} = r_{13}$
15	$E + CN \leftrightarrow ECN$	$k_{30} [CN][E] - k_{31} [ECN]$	(k30)30	(k31)30	$\frac{d[ECH]^*}{dt} = r_{14}$
16	$E + CN + F \rightarrow ECN + F$	$k_{32} [CN][E][F]$	(k32)40	-	$\frac{d[ECN]^*}{dt} = r_{15} + r_{16}$
17	$E + WN \leftrightarrow EWN$	$k_{21} [WN][E] - k_{22} [EWN]$	(k21)0.008	(k22)0.004	$\frac{d[EWN]^*}{dt} = r_{17} + r_{18} + r_{19}$
18	$E + WN + B \rightarrow EWN + B$	$k_{23} [WN][B][E]$	(k23)1	-	
19	$E + WN + 2 \cdot CN \leftarrow EWN + 2 \cdot CN$	$k_{25} [EWN][CN][CN]$	-	(k25)0.1	
20	$E + 3 \cdot T \leftrightarrow ET_3$	$k_{13} [T][T][T][E] - k_{14} [ET]$	(k13)100	(k14)26	

* Enzyme Substrate complex

Table 3 Simulation settings used for the uncertainty and sensitivity analysis

Uncertainty analysis	Sensitivity Analysis	
Monte Carlo simulations	Standardised Regression Coefficients	Morris Screening
Sampling input uncertainty 500 samples were selected. Parameters were considered to be uncorrelated due to unavailability of the information on the correlation matrix.	The R ² value for the linear regression for each of the model outputs of the Monte Carlo simulations should be greater than 0.7	The number of levels, p and number of repetitions, r were defined as 6 and 30, respectively
Input uncertainty of ± 50 % variability around the parameter estimates is used. Assumed uniform probability distribution		

imitated phase separations such as the equilibrium between W and WW.

3.2. Uncertainty and sensitivity analysis simulation settings

The settings used for the uncertainty and sensitivity analysis, are shown in Table 3. It should be noted that the approach by Sin *et al.* (2009) and Vangsgaard *et al.* (2012) in classifying the input factor uncertainty, is to class the parameters into three different groups, depending on their level of uncertainty. Given there are no other reported parameters for this reaction, the third class is chosen, which has the highest uncertainty of ± 50% variability around the parameter estimates.

3.3. Modelling and simulation environment

The model was implemented and simulated in Matlab. All the above mentioned methods for performing the uncertainty and sensitivity analysis were also implemented in Matlab as m script files building on the toolbox based on the work by Sin *et al.* (2009).

4. RESULTS AND DISCUSSION

4.1. Monte Carlo Simulations

The uncertainty in the model outputs is represented using the mean along with the 5th and 95th percentile of the distribution of each model output, obtained from the dynamic simulation of the 500 Latin hypercube samples. For the kinetic model investigated, only the typically measured variables (TAG, DAG, MAG, FFA, FFAE) during the transesterification reaction are reported.

Monte Carlo results: The Monte Carlo method was used to propagate the uncertainty of the kinetic parameters on the output (prediction) uncertainty of the model. The

interpretation of these results in Fig. 2 is straightforward; the higher the uncertainty band (95th and 5th percentiles), the worse the model prediction quality is.

The uncertainty analysis gives insight into the model structure. For example, the uncertainty on FFAE and TAG was relatively smaller compared to the uncertainty on the predictions of FFA, DAG and MAG. Furthermore, the uncertainty was observed to be changing over time during the reaction. These phenomena could be explained by the fact that as the reaction proceeds, the concentrations of the other components become more pronounced and the uncertainty of the model parameters affects the model outputs more. At the end of the reaction, the concentrations decrease, so the contribution from the parameters also decrease and hence cause a decrease in the uncertainty of the model outputs. This concept is reinforced if the FFAE concentration profile is investigated. For the ± 50 % variability in the parameter estimates, Fig. 2 gives a clear depiction of where the model is most reliable; in the start (first 5 hours) of the reaction.

The uncertainty analysis gives the modeller much more insight into how the model behaves compared to using a local sensitivity analysis method. Statistically speaking, uncertainty bands observed in Fig. 2 correspond to the distribution of the model outputs at each time instant. A look at the Cumulative distribution function paints a better picture of the acceptability of the model outputs. For example, in Fig. 2 the variances in the model outputs visually appear quite large after 10 hours. Choosing the time of 20 hours the Cumulative distribution function in Fig. 3 shows that the DAG concentration has a mean value of 0.07 mol/L with a standard deviation of 0.09 mol/L. This is a quite wide variation and depending on the application this may or may not be acceptable. Given the ± 50 % variability used on the parameter estimates, the decision maker now has statistical meaningful bounds on which to base further calculation. Take for example, an engineer, who wishes to do an economic evaluation on the final FFAE yield. At the reaction end time of 50 hrs the FFAE concentration has a mean value of 0.98 mol/L with a standard deviation of 0.02 mol/L

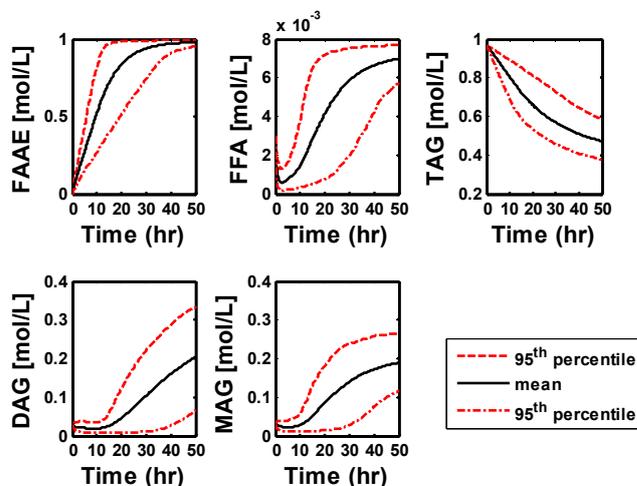


Fig. 2 Uncertainty analysis of the model predictions for FFAE, FFA, TAG, DAG, MAG (The mean and the 5th and 95th percentiles are obtained from performing 500 Monte Carlo simulations).

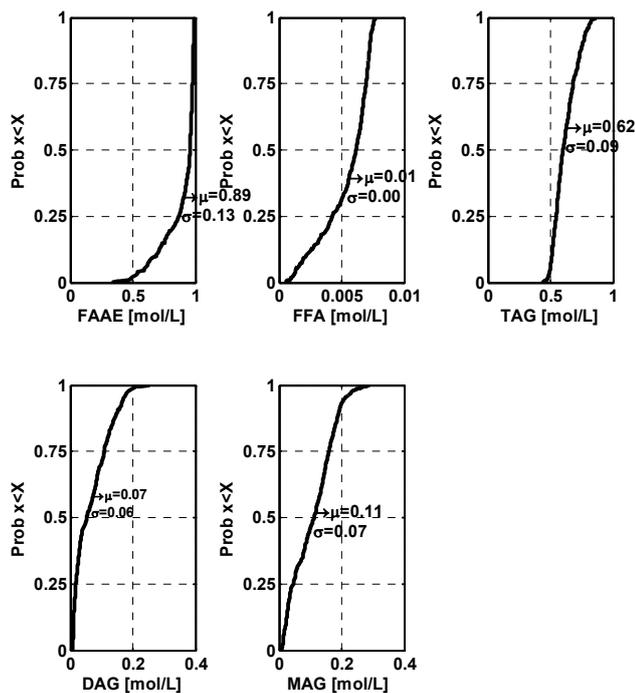


Fig. 3. Cumulative distribution function of the 5 model outputs at a time of 20 hrs. The x-axis shows the concentration of the model outputs and the y-axis shows the probability of a value in the x-axis being greater or equal to a chosen concentration. The mean along with the standard deviation for each model output is shown.

4.2. Sensitivity Analysis

For the calculation of the Standardised Regression Coefficients and the Morris screening a scalar output was needed. A meaningful property of time-series for each model output needs to be determined. The time of 20 hours is chosen for the analysis. This is done so as to compare the rankings of the parameters obtained, from the two sensitivity analysis methods.

Standard Regression Coefficients results: The degree of linearization indicated by the coefficient of model determination (R^2 value) obtained from the linear least squares fitting was over 0.8 for all model outputs. This indicates that the linearized model was able to explain most of the variance in the five model outputs investigated, and hence, the corresponding coefficients can reliably be used to assess the importance of the kinetic parameters on the model outputs. To allow for a detailed interpretation of the results, we restrict the further analysis to the parameters that have an absolute $\beta_{j,k}$ value greater than 0.1. This cut off value corresponds roughly to a 1% fraction of the total variance.

As can be seen from Fig. 4, only 10 of the 32 model parameters mainly contribute to the variance in the model outputs. The parameter which has a significant effect on all the model outputs for this mechanism is k1, which is the forward rate of converting TAG to DAG. This is reasonable, given the proposed transesterification mechanism is that TAG is converted to DAG, in the second step, DAG is converted to MAG and in the third step, MAG is converted to glycerol.

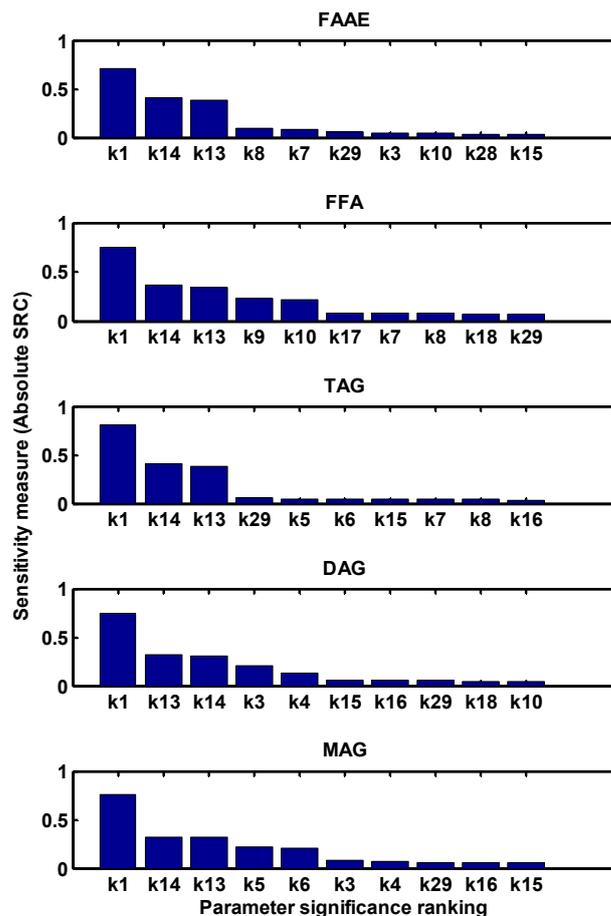


Fig. 4. The ranked absolute $\beta_{j,k}$ values (only the first 10 highest ranking parameters are shown) of the parameters influencing the transesterification reaction.

Also three FFA molecules would be produced in the process (this affects the FFA concentration) and the FFA is esterified to FFAE (This affects the FFAE concentration).

Most of the reaction rate constants, (k1-k12) are influential to the five model outputs, except for k2, k11 and k12. For example the parameters k9 and k10 which are for the hydrolysis of FFA, has considerable influence in the FFA model outputs. Of the binding and equilibrium constants (k13-k32) only k13 and k14 are sensitive. This means most of the binding parameters are non-influential. They can be fixed to any value within their ranges of uncertainty without significantly affecting the model prediction.

4.3. Morris Screening

To easily visualize the insignificant and significant parameters on the model output for the morris screening; the mean, μ_i and standard deviation, σ_i of the scaled elementary effects for each model output is plotted, along with the two lines formed from the standard error of the mean (1) (The standard error of the mean (SEM) gives an indication of the variability of the sample mean).

$$\mu_i = \pm 2.SEM = \pm 2 \frac{\sigma_i}{\sqrt{r}} \quad (1)$$

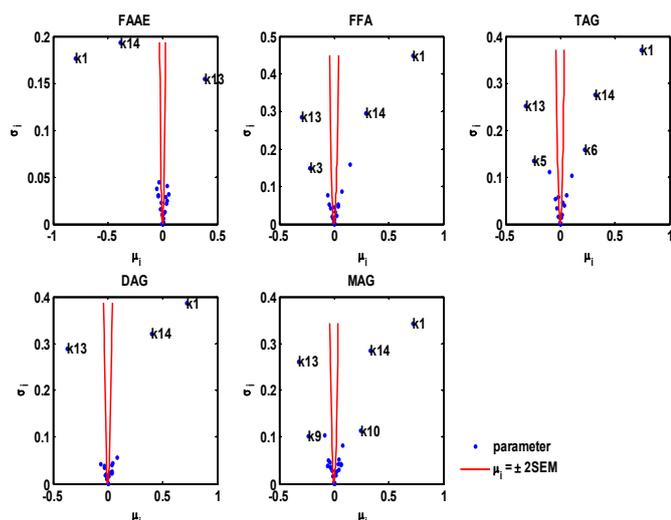


Fig. 5. Estimated mean and standard deviation of the distribution of elementary effects of the 32 parameters for the five model outputs. For clarity only the parameters with an extreme variation in μ_i are labelled.

These two lines form a wedge where if an input parameter lies inside the wedge, this indicates that its effect on the output is negligible and can be deemed insignificant. However, if the input parameter lies outside the wedge, then it is said to have a significant effect on the output.

Morris Screening results: The Morris Screening gives a good overview of the relative importance of uncertain factors as well as the associated non-linearity and interactions. For all of the model outputs the Morris Screening identifies a subset of factors classified as non-influential. The same parameters are identified as in the Standard Regression Coefficients method as being influential as seen in Fig 5. Parameters with the highest standard deviation, k1, k13 and k14 indicate the presence of non-linearity and/or interactions amongst the parameters (Campolongo *et al.* (2007)) (Note: A low standard deviation value, is an indication of a linear behaviour of the model for that particular parameter whereas a high standard deviation value indicates there may be nonlinear behaviour of the model for that particular parameter).

4.4. Comparison of Sensitivity analysis methods

It is noticed that both methods give different rankings for the parameters for both models. The question then arises, which tool should be used for further work? Given the resulting coefficient of determination, R^2 is greater than 0.7 for all the model outputs; this indicates that the Standard Regression Coefficients are a valid measure of sensitivity. The bonus is that the linear regression models could also be used in place of the original model within the linear model bounds. Also the values of the Standard Regression Coefficients hold physical meaning. The sign of the coefficient indicates the effect of the parameter on the model output. Example, for the FAME model output, k1 (forward rate constant for the conversion of TAG to DAG) has a positive Standard Regression Coefficients value of 0.71. An increase in the parameter estimate of k1 will cause an increase in FAME

production rate but a decrease in TAG production rate (TAG model output Standard Regression Coefficients is -0.80).

Morris Screening is found to give a good overview of the importance, interactions and non-linearity of the parameters. The method by Saltelli *et al.* (2007), used in this paper, considers both the mean and standard deviation of the scaled elementary effects, which makes the method more resilient to identifying a factor as influential when it is not (Type I errors). However, the method can be prone to Type II errors, that is, failing to identify a factor of considerable influence on the model (Saltelli *et al.* (2007)).

The two sensitivity analysis methods hence complement each other which were also found by Campolongo and Saltelli (1997).

4.5. Engineering Perspectives - Use of the kinetic models in enzymatic biodiesel simulation

The method used by Sin *et al.* (2009) and Vangsgaard *et al.* (2012) in classifying the parameter input uncertainty proves useful as a starting point to analyse how uncertainty in the parameters estimates influence the model outputs. If the confidence intervals for the parameter estimates were provided in the work by Fedosov *et al.* (2012) a more accurate evaluation of the parameter uncertainty on the model outputs could have been done. In general this highlights why when parameter estimates are being reported in literature, the confidence intervals should also be added. This will aid researchers in determining the reliability of the parameter estimates obtained along with being able to extend the uncertainty and sensitivity analysis with more rational bounds for the parameter uncertainty.

Uncertainty analysis was used to investigate the bounds and the applicability of a kinetic model for enzymatic biodiesel production. The distribution of the model outputs for the entire time course of the reaction are obtained from the uncertainty analysis. The engineer armed with knowledge of the model distribution can have greater confidence that the employed model, in the correct region, will obtain reproducible experimental results. For process development, an area of interest is reactor selection and configuration. A reliable kinetic model can be used to simulate and evaluate other reactor configurations, such as, fed-batch reactors and continuous stirred tank reactors (CSTR). The results from the simulations could then be verified in the lab. One hurdle to industrial implementation of enzymatic biodiesel production is enzyme inhibition by the alcohol substrate. Substrate feeding strategies can help mitigate these inhibition effects. Simulations can be used to devise an optimal feeding policy. For fed-batch operation, during the start or the reaction, the model investigated should perform adequately as seen from the Monte Carlo simulations. However one should be cautious. Take for example the situation where the model is used for simulating multiple CSTR in series. It may be problematic to correlate the experimental data, with what is simulated, if operating in regions of relatively high substrate and product concentrations. Given it is in this region the model has the most uncertainty.

In the field of process control the possibility of being able to make accurate predictions of all the states during the reaction enables the optimization of a feeding strategy for fed-batch operation (off line control). Another possibility is the combination of online measurements and the model to infer the evolution of the key reaction components i.e. a soft sensor or state observer. Information from the soft sensor is used as feedback to make controlled feeding of the substrate, optimizing the process performance. In order to design the observer it is necessary to have a process model and an estimate of the noise contribution from both the model and the online measurements. The uncertainty analysis in this case provides an estimate of the noise contribution from the model due to the uncertainty in the parameter estimates. One example is in the case of using viscosity measurements. In order to monitor the progress of the transesterification reaction, Ellis *et al.* (2008) used an in-situ viscometer to correlate the viscosity measurement with the rate of biodiesel production. To extend the work done by Ellis *et al.* (2008), the rate of biodiesel production, can be coupled with the kinetic model, which can then be used to infer the concentration of the five main components measured. Use of the viscosity meter along with the soft sensor provides a relatively inexpensive way to have real time monitoring of the system; giving to the operator quick feedback on the progress of the reaction and on the activity of the enzymes.

For predictive purposes such as determining when the transesterification is complete and tracking of the FFAE production the results from the uncertainty analysis showed that the parameter estimates has great potential; given the tight confidence intervals in the areas of interest. It should be noted that given the reaction mixture changes during the reaction (e.g. viscosity), most likely the rate constants also change during the reaction. Hence the rate constants found during the parameter estimation are just average values. This makes the uncertainty analysis a powerful tool to cater for the assumption that the parameters are fixed throughout the reaction.

In this study, the focus of the sensitivity and uncertainty analysis were on the parameters of the kinetic model. The uncertainty and sensitivity analysis tools could also be used to study the effects of the process conditions. It is known the different components in the oil (TAG, DAG, FFA etc.) can vary in concentration, and this could also be investigated to ascertain the effects on the model outputs. Sensitivity and uncertainty analysis on the process conditions can be an invaluable tool for the engineer in devising strategies to mitigate changes in the plant outputs due to uncertainties in the plant process conditions.

To get a clearer understanding of which kinetic parameters are responsible for the output uncertainty, two sensitivity methods (Standardized Regression Coefficients and Morris screening) help to distinguish the most influential parameters in the model. The influential parameters help elucidate which parameters contribute the most to the overall model outputs. For example the methanolysis parameters (k_7 and k_8) are influential for the FFAE model output while the hydrolysis parameters (k_9 and k_{10}) are influential for the FFA output.

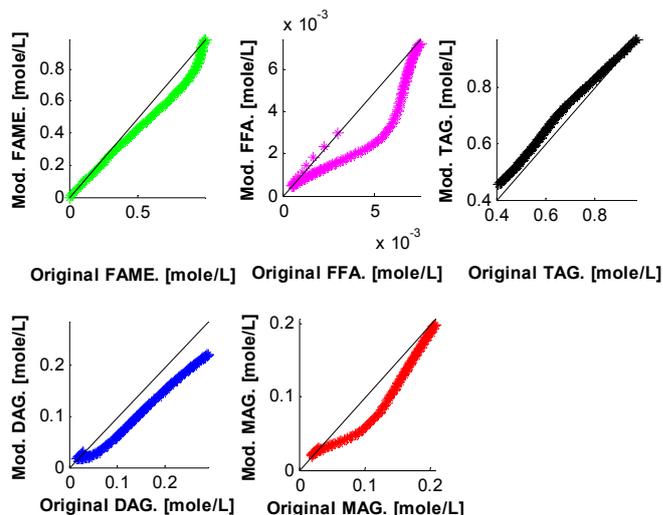


Fig. 6. Plot showing the Modified model prediction vs Original model predictions; where for the modified model predictions the non-influential parameters are removed. A R^2 value of over 0.92 is obtained for all plots.

The sensitivity analysis helps ascertain which mechanism dominates for a particular reaction. The non-influential parameters show where research needs to be placed in devising experiments to estimate those parameters. This is quite important if recalibrating the model parameters for a different enzyme or type of substrate. For the case presented by Fedosov *et al.* (2012), to aid in the model simplification, the non-influential parameters (k_2 , k_{11} , k_{12} and k_{15-32}) were removed. The parity plots are shown in Fig. 6, and show that the parameters removed, are essentially non-influential over the entire range of the reaction.

It should be noted, in the case of possible model reduction, care should be taken in removing parameters. If there is non-linearity or interactions amongst the parameters, a parameter with low importance according to ranking does not necessarily imply the factor to be non-influential. Therefore it is better to fix the value of the parameter, which is the recommendation for the equilibrium constants when recalibrating the model parameters for a different enzyme or type of substrate. These ideas brought out in this section are represented in the methodology outcomes branch in Fig. 1.

5. CONCLUSIONS

A kinetic model describing the enzymatic transesterification of vegetable oil was investigated, using Monte Carlo simulations of the model outputs, along with two sensitivity analysis methods based on screening and regression. The main points gleaned are:

1. The Monte Carlo simulations on the parameter estimates highlight the regions where the model is most reliable under uncertainty in the parameter estimates; in the start (first 5 hours) of the reaction. It is in this region that there is a greater probability of correlating the experimental data with what is simulated. This is an important for model based process development given the need for the simulation to represent reality. Also the 5th and 95th percentile of the distribution of each model output can be used in model-based

decision making such as bounds for economic process evaluation.

2. The sensitivity analysis successfully detected the influential and non-influential parameters to the model outputs. This “sets the stage” for model simplification when recalibrating the model parameters for different substrates. The non-influential parameters can be fixed at any given value within their ranges of uncertainty without affecting significantly the model output.

It should also be noted that the approach used in this paper is not only specific to the case of enzymatic biodiesel production but for other process as well.

6. ACKNOWLEDGMENTS

This work was partly supported by the Danish National Advanced Technology Foundation (DNATF) whose support is gratefully acknowledged.

7. REFERENCES

- Al-Zuhair, S. (2005) Production of Biodiesel by Lipase-Catalyzed Transesterification of Vegetable Oils: A Kinetics Study. *Biotechnol Progress*, 21(5), 1442–1448.
- Calabrò, V., Ricca, E., Paola, M. de, Curcio, S., and Iorio, G. (2010) Kinetics of enzymatic trans-esterification of glycerides for biodiesel production. *Bioprocess Biosyst Eng*, 33(6), 701–710.
- Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. *Environ Model Softw* 22(10): 1509–1518
- Campolongo, F. and Saltelli, A. (1997) Sensitivity analysis of an environmental model: an application of different analysis methods. *Reliab Eng Syst Saf*, 57(1), 49–69.
- Cheirsilp, B., H-Kittikun, A., and Limkatanyu, S. (2008) Impact of transesterification mechanisms on the kinetic modeling of biodiesel production by immobilized lipase. *Biochem Eng J*, 42(3), 261–269.
- Ellis, N., Guan, F., Chen, T., and Poon, C. (2008) Monitoring biodiesel production (transesterification) using in situ viscometer. *Chem Eng J*, 138(1-3), 200–206.
- Fedosov, S. N., Brask, J., Pedersen, A. K., Nordblad, M., Woodley, J. M., and Xu, X. (2012) Kinetic model of biodiesel production using immobilized lipase *Candida antarctica* lipase B. *J Mol Cat B: Enzy0*, 85-86, 156–168.
- Foss, B., Lohmann, B., and Marquardt, W. (1998) A field study of the industrial modeling process. *J Process Control*, 8(5-6), 325–338.
- Heitzig, M., Sin, G., Sales-Cruz, M., Glarborg, P., and Gani, R. (2011) Computer-Aided Modeling Framework for Efficient Model Development, Analysis, and Identification: Combustion and Reactor Modeling. *Ind Eng Chem Res*, 50(9), 5253–5265.
- Li, W., Li, R.-w., Li, Q., Du, W., and Liu, D. (2010) Acyl migration and kinetics study of 1(3)-positional specific lipase of *Rhizopus oryzae*-catalyzed methanolysis of triglyceride for biodiesel production. *Process Biochem*, 45(12), 1888–1893.
- Lv, D., Du, W., Zhang, G., and Liu, D. (2010) Mechanism study on NS81006-mediated methanolysis of triglyceride in oil/water biphasic system for biodiesel production. *Process Biochem*, 45(4), 446–450.
- Pilarek, M. and Szweczyk, K. W. (2007) Kinetic model of 1,3-specific triacylglycerols alcoholysis catalyzed by lipases. *J Biotechnol*, 127(4), 736–744.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2007) *Global Sensitivity Analysis. The Primer*, John Wiley & Sons, Ltd, Chichester, UK.
- Saltelli, A., Ratto, M., Tarantola, S., and Campolongo, F. (2006) Sensitivity analysis practices: Strategies for model-based inference. *Reliab Eng Syst Saf*, 91(10-11), 1109–1125.
- Sin, G., Gernaey, K. V., and Lantz, A. E. (2009) Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol Progress*, 25(4), 1043–1053.
- Vangsgaard, A. K., Mauricio-Iglesias, M., Gernaey, K. V., Smets, B. F., and Sin, G. (2012) Sensitivity analysis of autotrophic N removal by a granule based bioreactor: Influence of mass transfer versus microbial kinetics. *Bioresour Technol*, 123, 230–241.