IMPLEMENTATION OF MULTILINEAR LOOK-UP TABLES AS SURROGATE MODELS FOR THE ETHYL TRANSESTERIFICATION REACTOR IN EQUATION-ORIENTED SIMULATOR EMSO

IMPLEMENTAÇÃO DE INTERPOLADORES MULTILINEARES COMO MODELOS SIMPLIFICADOS DE REATORES DE TRANSESTERIFICAÇÃO ETANÓLICA NO SIMULADOR GLOBAL EMSO

C. O. MARTINS[1, *], F. F. FURLAN[1] and R. C. GIORDANO[1]

[1] Federal University of São Carlos, Department of Chemical Engineering, São Carlos, São Paulo, Brazil

*Corresponding author. Federal University of São Carlos, Department of Chemical Engineering, São Carlos, São Paulo, Brazil, Phone: +55 16 991941404
e-mail address: christiandeoliveiramartins@gmail.com (C. O. Martins).

A B S T R A C T

The techno-economic analysis of chemical processes is usually based on steady state simulations. When dynamic processes are present, simplified or surrogate models are used to avoid the dynamic simulation of the overall system. Nevertheless, this usually results in loss of information when the simplified model is used or a large number of parameters when the surrogate model is applied. To circumvent this problem, a hybrid surrogate model is proposed that combines the low-cost prediction of the average behavior by a simplified model with the fine tuning provided by a multilinear look-up table. This concept was applied to the homogeneous alkaline transesterification of the soybean oil with ethanol in a batch reactor. The hybrid approach enabled a 70% reduction in the number of points of the look-up table compared to the pure interpolator model, for an accuracy tolerance of 0.01 mol/l.

R E S U M O

A análise técnico-econômica de processos químicos costuma se basear em simulações estacionárias. Quando processos dinâmicos estão presentes, modelos simplificados ou meta-modelos são usados para evitar a simulação dinâmica de todo processo. Mas isso normalmente resulta em perda de informação, no caso do modelo simplificado, ou em um grande número de parâmetros no caso dos meta-modelos. Para contornar isto, um meta-modelo hibrido é proposto, que combina a predição de baixo custo computacional do comportamento médio do sistema por um modelo simplificado, com o refinamento provido por um interpolador multilinear. Esta metodologia foi aplicada na transesterificação alcalina homogênea do óleo de soja com etanol em um reator batelada. O modelo híbrido reduziu em 70% o número de pontos da tabela de inspeção para uma tolerância de 0.01 mol/l de precisão comparado ao modelo de interpolador puro.

. 1. INTRODUCTION

It is already a consensus that a shift towards a carbon neutral economy is mandatory. To allow this transition, multiple options and sources will probably coexist, with biomass as an important source for both energy and materials (Giordano et al., 2020). In this context, a good alternative for petrochemical derived fuels is biodiesel. Biodiesel is a monoalkyl ester of a fatty acid produced from renewable sources, such as vegetable oils and animal fat (Mahlia et al., 2020).

The biodiesel production is usually based on the transesterification reaction of a triacyl glyceride with an alcohol. It can be homogeneously catalyzed by acids, bases and enzymes or heterogeneously by oxides of alkaline earth metals, zeolites, immobilized enzymes, among other options (Guerrero, 2011). Methanol is mainly used as the alcoholic compound. Nevertheless, there is a trend in Brazil towards the use of ethanol, given the potential of the ethanol and sugar industry and the possibility to achieve a more renewable fuel through this route (Guerrero, 2011).

Most of the catalysts options listed are still not consolidated in industry. Therefore, it is of paramount importance to perform techno-economic and environmental analyses to assist their development towards industrial application (Furlan et al., 2016). Since these analyses are usually based on steady-state simulations, essentially dynamic processes, such as batch and feed batch reactors, are mainly represented by simplified models (stoichiometric reactors, for example). These simplified models present low accuracy for describing the influence of process conditions on the system performance.

An option to circumvent this problem is to use surrogate models. Surrogate models aim at representing complex and costly systems using cheaper functions. Several surrogate modeling functions are available (Alizadeh et al., 2020). One option is the multilinear look-up table (Nelles, 2001), which consists of a set of data points positioned in a multi-dimensional grid. These data are based on the simulation of the rigorous model. The surrogate output is obtained by the interpolation of the points of the smallest n-dimensional cube that contains the point of interest. The advantage of this type of surrogate model is the lack of a training step, since the data is fully used. A disadvantage is that it demands a large number of points to describe the output of nonlinear systems accurately.

In this context, this study presents a hybrid model composed by an analytic solution of an approximation of the original ODE and a multilinear look-up table. The former is responsible for a rough approximation of the output while the latter improves the local behavior. The hybrid model is capable of accurately representing the dynamic model behavior with a smaller number of parameters, when compared to the pure multilinear look-up table. The approach was applied to the ethylic transesterification of soybean oil. The simulation was performed in EMSO (Soares & Secchi, 2013), an equation-oriented process simulator.

2. MATERIALS AND METHODS

All the simulations were performed in EMSO, which is an equation-oriented process simulator with an internal object-oriented modelling language (Soares & Secchi, 2003). EMSO is capable of performing process simulations, optimizations, parameter estimation and data reconciliation. It has an interface with MATLAB through Simulink and python. The look-up table model, used in this study, was included in EMSO via a plug in (Furlan et al., 2016), since it is not well suited for the equation-oriented approach due to its algorithmic modelling.

The rigorous model for the transesterification of soybean oil with ethanol, using sodium ethoxide as catalyst, is presented by Dias (2016). The global reaction rates are based on the pseudo-components that represent the triacyl glycerides (TG), diacyl glycerides (DG) and monoacyl glycerides (MG), the ethyl esters that compose the biodiesel (EE), along with the glycerol (GL) and the ethanol (ET). This reaction occurs in three steps, represented by reversible reactions described by Equations 1-3.

$$TG + ET \rightleftarrows DG + EE \tag{1}$$

$$DG + ET \rightleftarrows MG + EE \tag{2}$$

$$MG + ET \rightleftarrows GL + EE \tag{3}$$

The kinetics models of these reactions are represented by Equations 4-9.

$$r_{TG} = -k_1 \cdot C_{TG} \cdot C_{ET} + k_2 \cdot C_{DG} \cdot C_{EE} \tag{4}$$

$$r_{DG} = k_1 \cdot C_{TG} \cdot C_{ET} - k_2 \cdot C_{DG} \cdot C_{EE} - k_3 \cdot C_{DG} \cdot C_{ET} + k_4 \cdot C_{MG} \cdot C_{EE} \tag{5}$$

$$r_{MG} = k_3 \cdot C_{DG} \cdot C_{ET} - k_4 \cdot C_{MG} \cdot C_{EE} - k_5 \cdot C_{MG} \cdot C_{ET} + k_6 \cdot C_{GL} \cdot C_{EE} \tag{6}$$

$$r_{EE} = k_1 \cdot C_{TG} \cdot C_{ET} - k_2 \cdot C_{DG} \cdot C_{EE} + k_3 \cdot C_{DG} \cdot C_{ET} - k_4 \cdot C_{MG} \cdot C_{EE} + k_5 \cdot C_{MG} \cdot C_{ET} - k_6 \cdot C_{GL} \cdot C_{EE} \tag{7}$$

$$r_{GL} = k_5 \cdot C_{MG} \cdot C_{ET} - k_6 \cdot C_{GL} \cdot C_{EE} \tag{8}$$

$$r_{ET} = -r_{EE} \tag{9}$$

Where r_i is the global reaction rate of each compound 'i' (TG, DG, MG, EE, GL, ET)., k_n is the kinetic constant of each specific reaction "n" and C_i is the concentration of each compound "i".

Two approaches are compared: a pure interpolator that directly approximates the rigorous model's output using a look-up table; and a hybrid model, composed by a simplified model of exponential equations which represents a rough approximation of the rigorous model combined to an interpolator that refines the output. The pure interpolator demands many points to make the look-up table. Therefore, the latter is proposed as a way to reduce the number of points necessary to produce an interpolator with the desired precision. For this purpose, equations (4) to (9) were modified so they formed a system of equations with known algebraic solution that still approximates the rigorous model output behavior, representing the dynamic nature of the system. Ergo the interpolator is just responsible for the correction of the deviation between the rigorous model and this simplified model, which presents a non-linear behavior yet, but with less variations.

The considerations made to provide the simplified model were:

- The concentration of ethanol is considered constant during the reaction;
- The kinetics of the reverse reactions are disregarded; so, the solutions for the kinetics

of the reactions of formation and consumption of the acyl glycerides are defined analytically;

- The other compounds are defined by stoichiometry.

These conditions create a system of ordinary differential equations represented by Equations 10-17.

$$r_{TG} = -k_1' \cdot C_{TG} \tag{10}$$

$$r_{DG} = k_1' \cdot C_{TG} - k_3' \cdot C_{DG} \tag{11}$$

$$r_{MG} = k_3' \cdot C_{DG} - k_5' \cdot C_{MG} \tag{12}$$

$$r_{EE} = k_1' \cdot C_{TG} + k_3' \cdot C_{DG} + k_5' \cdot C_{MG} \tag{13}$$

$$r_{GL} = k_5' \cdot C_{MG} \tag{14}$$

$$r_{ET} = -r_{EE} \tag{15}$$

$$k_n' = kk_n \cdot C_{ETo} \tag{16}$$

Where k_n' represents the product of the pseudo-kinetic constants kk_n, analogue to the real kinetic constant, and the initial concentration of ethanol C_{ETo}. The batch reactor molar balance is given by Equation 16.

$$\frac{dC_i}{dt} = r_i \tag{17}$$

Where t represents the time of reaction.

Then, we can achieve an analytic solution to describe the concentration of the acyl glycerides with the Equations 18-20.

$$C_{TG} = C_{TGo} \cdot e^{-k_1' \cdot t} \tag{18}$$

$$C_{DG} = \frac{k_1' \cdot C_{TGo}}{k_3' - k_1'} \cdot \left(e^{-k_1' \cdot t} - e^{-k_3' \cdot t} \right) + C_{DGo} \cdot e^{-k_3' \cdot t} \tag{19}$$

$$C_{MG} = k_3' \cdot \left[\frac{k_1' \cdot C_{TGo}}{k_3' - k_1'} \cdot \left(\frac{e^{-k_1' \cdot t} - e^{-k_3' \cdot t}}{k_5' - k_1'} - \frac{e^{-k_3' \cdot t} - e^{-k_5' \cdot t}}{k_5' - k_3'} \right) + \frac{C_{DGo} \cdot \left(e^{-k_5' \cdot t} - e^{-k_5' \cdot t} \right)}{k_5' - k_3'} \right] + C_{MGo} \cdot e^{k_5' \cdot t} \tag{20}$$

These new constants were fit to the rigorous model data. Model fitting was performed using "Parameter Estimation" environment in software EMSO (Soares, 2007). Since the new system of equations presents just three independent reactions, it was only necessary to obtain the concentrations of the acyl glycerides by the rigorous model. The other concentrations were obtained using the reactions stoichiometries and mass balances. With these manipulations of Equations 18-20, the concentrations of ethanol, ethyl esters (biodiesel) and glycerin are represented by Equations 21-23.

$$C_{ET} = C_{ETo} + 3 \cdot (C_{TG} - C_{TGo}) + 2 \cdot (C_{DG} - C_{DGo}) + C_{MG} - C_{MGo} \tag{21}$$

$$C_{EE} = -3 \cdot (C_{TG} - C_{TGo}) - 2 \cdot (C_{DG} - C_{DGo}) - (C_{MG} - C_{MGo}) \tag{22}$$

$$C_{GL} = C_{MGo} - C_{MG} + C_{DGo} - C_{DG} + C_{MGo} - C_{MG} \tag{23}$$

The inputs of the multilinear look-up table based surrogate model were the reaction time and the molar ratio between ethanol and soybean oil at the beginning of the reaction. The concentration of soybean oil used was 700 mol/m³, the same value used during the experiments from the reference (Dias, 2016).

In the base case, the surrogate model was responsible for approximating the behavior of the rigorous model. On the other hand, in the hybrid surrogate model, the interpolator was responsible for correcting the deviation between the outputs of the rigorous model and the simplified model (Equation 24). Therefore, the output of the hybrid surrogate model is the sum of the simplified model one and the interpolator one (Equation 11). A MATLAB® based program developed by Lino (2018) was used to construct the look-up table using its interface with EMSO. The program systematically refines the table until the max absolute error (MAE - Equation 26) achieves a specified tolerance. A Latin-hypercube sampling approach was used for model validation (Saltelli et al., 2008) with 1000 validation points.

$$d_i = C_i^{rm} - C_i^{sm} \tag{24}$$

$$C_i^{hm} = C_i^{sm} + d_i \tag{25}$$

$$MAE = max\left(\left| C_i^{hm} - C_i^{rm} \right| \right) \tag{26}$$

Where d_i is the deviation from the rigorous model concentration (C_i^{rm}), C_i^{sm} is the simplified model concentration, C_i^{hm} is the hybrid surrogate model concentration and MAE is the max absolute error.

In the pure look-up table case, the MAE is calculated with the concentration from the look-up table instead of the concentration from the hybrid model.

3. RESULTS AND DISCUSSION

Initially, the simplified model was fit to the rigorous model outputs. The used data were generated by the simulation of the rigorous model. The output variables were the concentrations of TG, DG and MG. The input variables were the time, evaluated at values between 0 and 25 minutes with 0.5 minutes steps, and the ethanol/soybean oil molar ratio, evaluated at values between 6 and 12 with 0.5 steps. The parameters values obtained are shown in Table 1.

Table 1 – Results of simplified model's pseudo constant fitting.

Pseudo-constant	Value / $m^3 \cdot mol^{-1} \cdot s^{-1}$	R^2 / #
kk_1	1.295×10^{-6}	0.99
kk_2	8.717×10^{-7}	0.83
kk_3	8.700×10^{-7}	0.85

The deviation between the behavior of the simplified model and the rigorous model are shown in Figure 1. The concentrations calculated by the simplified model and the rigorous model are shown in Figure 2, both for an initial ethanol/soybean oil molar ratio of 9/1 and 700 mol/m³ of initial concentration of soybean oil, for the mono (MG), di (DG) and

triacyl glycerides (TG). As it can be seen in both figures, the simplified model systematically under and over estimates the concentrations depending on the region.
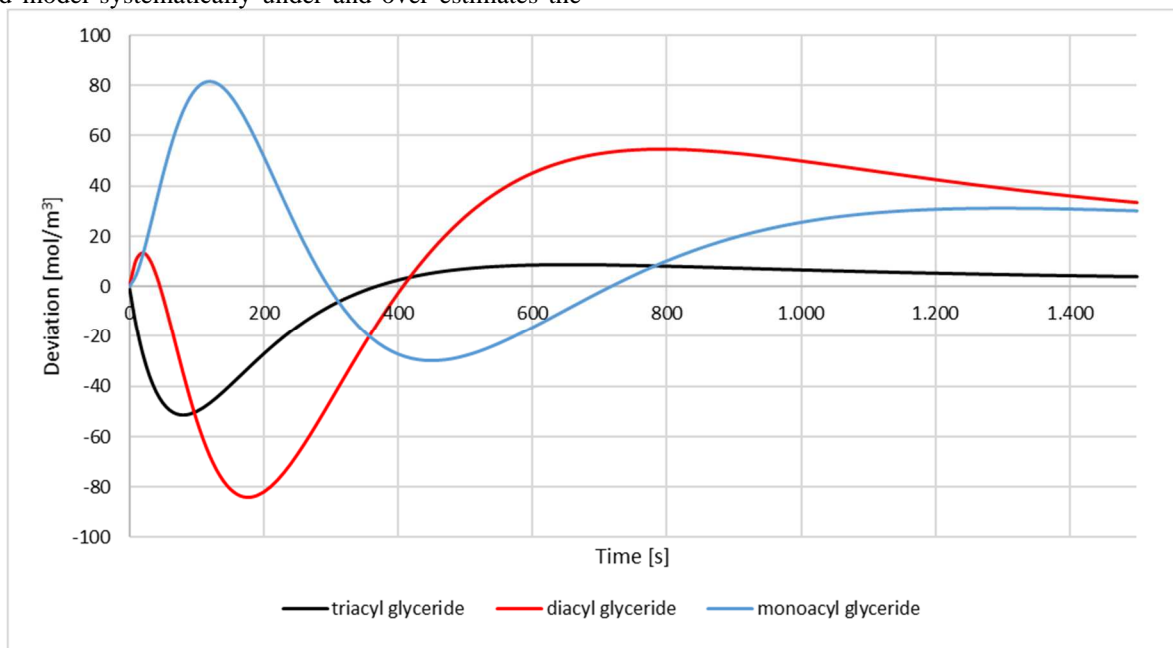

Figure 1 – Deviation, d_i, between the concentrations calculated by the rigorous model and the simplified model for each acyl glyceride.
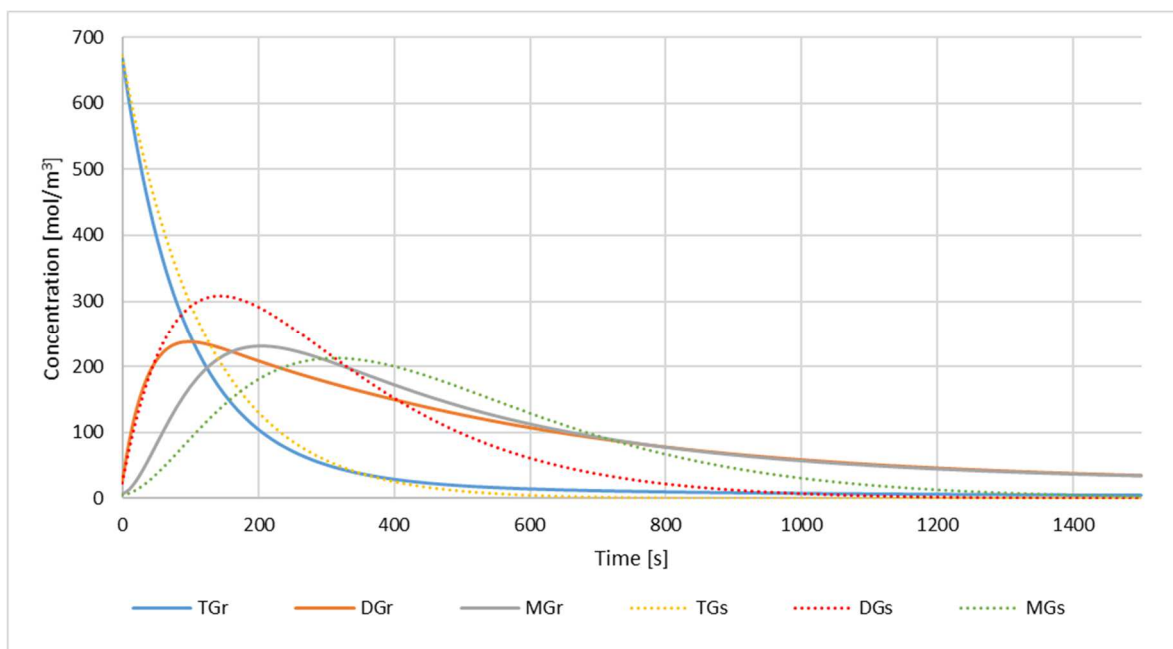

Figure 2 – Concentrations of TG, DG and MG calculated by the rigorous model ("r" index) and by the simplified model ("s" index).

For the construction of the hybrid surrogate model, that unites the simplified model and the multilinear look-up table, the maximum absolute error of the final model was considered as stopping criteria. The grid of points used by the interpolator was chosen as the one which achieve the specified accuracy tolerance in the validation test consisting of 1000 points. In Figure 3 the correspondence between the results of the simplified model and the rigorous model are shown for the diacyl glyceride (DG), the compound with the highest values of error. The results are at an ethanol/soybean oil molar ratio equal to 9 and initial concentration of soybean oil equal 700 mol/m³. Figures 4 and 5 shows the comparison between the rigorous model and the hybrid model for the tolerances of 100 and 10 mol/m³, respectively. As expected, as the maximum absolute error allowed decreases, the systematic error presented by the simplified model is reduced. The improvement is also noticed by the increase in the coefficient of determination value (R^2). The pure simplified model presented a R^2 of 0.851. The hybrid model presented R^2 of 0.919 and 1.00, for tolerances of 100 and 10 mol/m³, respectively.
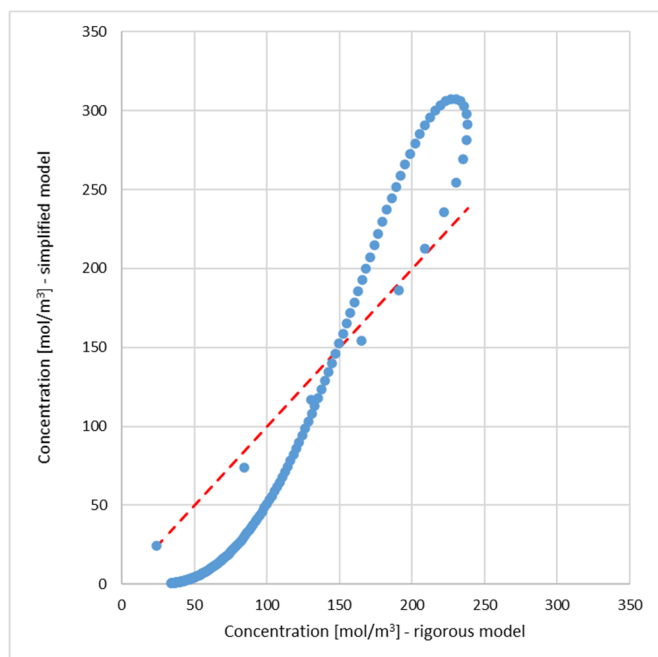
Figure 3 – DG concentration obtained from rigorous dynamic model versus the concentration predicted by the simplified model.
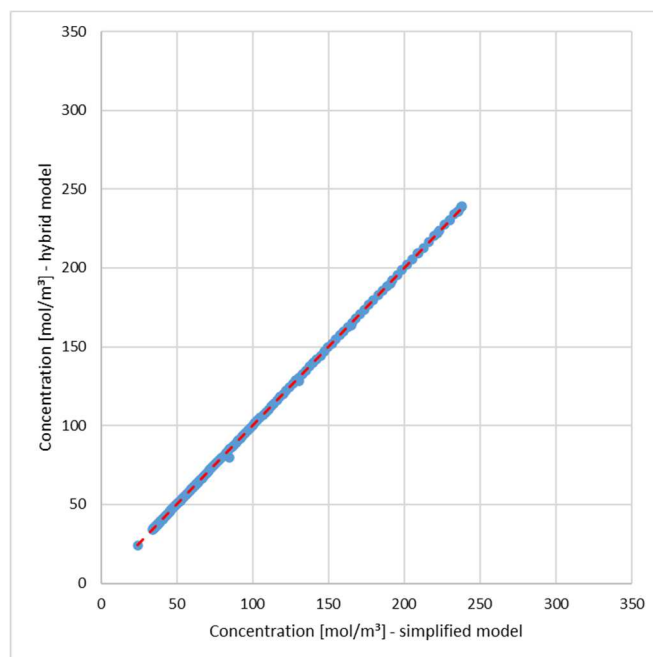

Figure 5 – DG concentration obtained from rigorous dynamic model versus the hybrid model with 10 mol/m³ error tolerance.
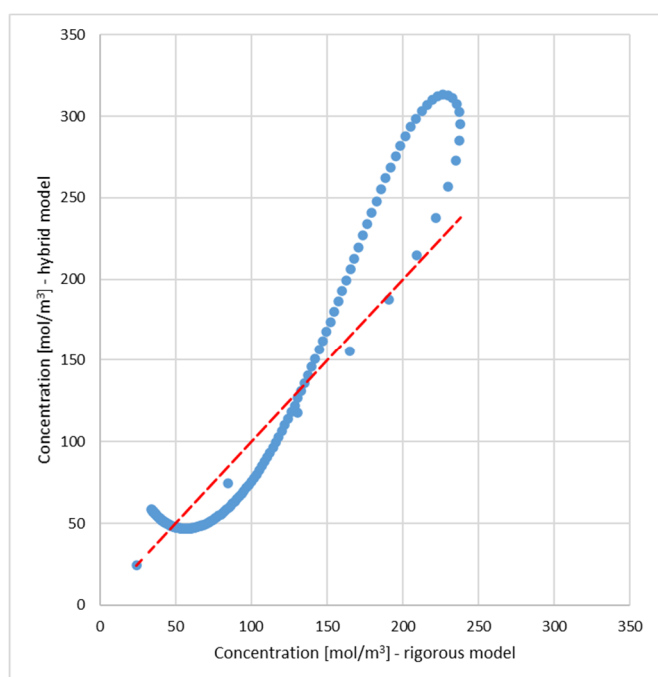

Figure 4 – DG concentration obtained from rigorous dynamic model versus the concentration predicted by the hybrid model with 100 mol/m³ error tolerance.

Finally, the hybrid surrogate model was compared to the pure interpolator surrogate model. For this, the number of points necessary for the look-up table to achieve certain precision of each model was observed. Table 2 and Table 3 present some results of the number of points and the average evaluation time of ten simulations varying the ethanol/soybean oil molar ratio by 6.6 to 12 with steps of 0.6. It can be seen that the number of points in the hybrid model was 70-88% less than the pure interpolator model for the same tolerance. Finally, this is translated in a decrease in evaluation time, with the hybrid model showing average evaluation times 40-52% smaller than the pure interpolator model. It can be seen that the rigorous model evaluation time is 1.65 times the hybrid model's for the 10 mol/m³ threshold. Although this difference is apparently small, it is expected to increase when the whole process (including upstream and downstream) is considered. The simulations were made in a computer with a Core i3-4005U 1.7GHz and 3M cache. The time of simulation of the rigorous model for the same ten simulations was 1.44±0.05.

Table 2 – Number of points and simulation time for the hybrid surrogate model and the pure interpolator surrogate model for an error tolerance of 100 mol/m³.

Analysis parameters	Pure look-up table model	Hybrid model	Reduction / %
Number of points / #	34	4	88
Simulation time / s	0.30±0.02	0.18±0.06	40

Table 3 – Number of points and simulation time for the hybrid surrogate model and the pure interpolator surrogate model for an error tolerance of 10 mol/m³.

Analysis parameters	Pure look-up table	Hybrid model	Reduction / %
Number of points / #	645	195	70
Simulation time / s	1.82±0.10	0.88±0.09	52

It can be noticed that, as the error tolerance is reduced, the difference in the number of points of the hybrid and the pure look-up table models decreases. This indicates that the impact of the approximation by the simplified model in the reduction of the number of points of the look-up table decreases. On the other hand, the difference in the simulation time of the two models increases, which indicates that the increase in the look-up table evaluation time is superlinear with the number of points. Also, the simulation time of the hybrid model was faster than the rigorous model for both tolerances tested, while the pure look-up table presented a higher simulation time for the 10 mol/m³ tolerance.

Through this approach it is possible to integrate batch process models with continuous models without loss of information. For the present case study, this will allow the simulation, techno-economic analysis, and optimization of the global flowsheet of an ethyl biodiesel production, which includes a continuous purification step with distillation columns, evaporators and decanters (Marchetti, Miguel e Errazu, 2008; Santana et al., 2010; Campos, 2009).

4. CONCLUSION

The use of multilinear look-up tables as surrogate models of dynamic systems in steady-state simulations leads to a high number of parameters. The hybrid surrogate model proposed reduces the complexity of the model using a linear approximation of the ordinary differential equations, with known analytic solution. Using this method, the number of necessary points for the surrogate model to achieve certain accuracy was reduced between 70 and 88%, depending of the chosen accuracy. Additionally, the simulation time had a reduction of 40-52%, and the simulation time of the hybrid model was minor than the rigorous model, on the contrary of pure multilinear look-up table model. Although the surrogate model was used in the case study of homogeneous basic transesterification of soybean oil, the approach is general and can be applied to any system of differential equation. This method ensures the possibility future analysis of the use of batch reactor in global flowsheets of production of biodiesel.

5. ACKNOWLEDGES

R E F E R E N C E S

ALIZADEH, R.; ALLEN, J. K.; MISTREE, F. MANAGING COMPUTATIONAL COMPLEXITY USING SURROGATE MODELS: A CRITICAL REVIEW. RESEARCH IN ENGINEERING DESIGN, V. 31, N. 3, P. 275–298, 2020.

CAMPOS, R. J. A. METODOLOGIA DE ANÁLISE ENERGÉTICA E EXERGÉTICA APLICADA À PLANTA PILOTO DE BIODIESEL DE CAETÉS-PE. 2009. 137f. Dissertação (Mestrado em Engenharia Mecânica) – Programa de Pós-Graduação em Engenharia Mecânica, Universidade Federal de Pernambuco, Recife, 2009. Available in: https://repositorio.ufpe.br/bitstream/123456789/5123/1/arquivo2141_1.pdf. Accessed in: August 2020.

DIAS, T.P.V.B., ESTUDO DO EQUILÍBRIO LÍQUIDO–LÍQUIDO E CINÉTICA DE REAÇÃO NA PRODUÇÃO DE ÉSTERES ETÍLICOS. 2016. 171f. Tese (Doutorado em Engenharia de Alimentos) – Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas, Campinas, 2016. Available in: http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/255100/1/Dias_TelmaPorcinaVilasBoas_D.pdf. Accessed in: July 2020.

FURLAN, F. F.; COSTA, C. B. B.; SECCHI, A. R.; WOODLEY, J. M.; GIORDANO, R. C. Retro-techno-economic analysis: using (bio) process systems engineering tools to attain process target values. Industrial & Engineering Chemistry Research, v. 55, p. 9865-9872, 2016a.

FURLAN, F. F.; LINO, A. R. A.; MATUGI, K.; CRUZ, A. J. G.; SECCHI, A. R.; GIORDANO, R. C. A simple approach to improve the robustness of equation-oriented simulators: Multilinear look-up table interpolators. Computers and Chemical Engineering, v. 86, p. 1-4, 2016b.

GIORDANO, R. C.; ELIAS, A. M.; FURLAN, F. F. Virtual Biorefineries: computational tools for bridging the gap toward a low carbon economy. In: Filho, E. X. F. et al. (Eds). Recent Advances in Bioconversion of Lignocellulose to Biofuels and Value-added Chemical Within the Biorefinery Concept. Elsevier, 2020, p. 233-257.

GUERRERO, J.K.R.G., SÍNTESE DE BIODIESEL A PARTIR DE ÓLEO DE MAMONA EMPREGANDO ETANOL SUBCRÍTICO. 2011. 100f. Dissertação (Mestrado em Engenharia Química) – Faculdade de Engenharia Química, Universidade Estadual de Campinas, Campinas, 2011. Available in: http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/266891/1/RodriguezGuerrero_JulieKaterine_M.pdf. Accessed in: July 2020.

LINO, A. R. A., Interpoladores multilineares como metamodelos: Integração à simulação de biorrefinarias no ambiente EMSO. 2018. 131f. Dissertação (Mestrado em Engenharia Química) – Departamento de Engenharia Química, Universidade Federal de São Carlos, São Carlos,2018. Available in: https://repositorio.ufscar.br/bitstream/handle/ufscar/10213/LINO_Anderson_2018.pdf?sequence=4&isAllowed=y. Accessed in: July 2020.

MAHLIA, T. M. I. et al. Patent landscape review on biodiesel production: Technology updates. **Renewable and Sustainable Energy Reviews**, v. 118, n. October 2019, p. 109526, 2020.

MARCHETTI, J. M.; MIGUEL, V. U.; ERRAZU, A. F.

Techno-economic study of different alternatives for biodiesel production. Fuel Processing Technology, v. 89, n. 8, p. 740–748, 2008.

NELLES, O. Nonlinear System Identification. [s.l.] Springer Berlin Heidelberg, 2001. MARCHETTI, J. M.; MIGUEL, V. U.; ERRAZU, A. F. Techno-economic study of different alternatives for biodiesel production. Fuel Processing Technology, v. 89, n. 8, p. 740–748, 2008.

SALTELLI, A.; RATTO, M.; ANDRES, T.; CAMPOLONGO, F.; CARIBONI, J.; GATELLI, D.; SAISANA, M.; TARANTOLA, S. Global Sensitivity Analysis. The Primer. [s.l: s.n.]. v. 304

SANTANA, G. C. S. et al. Simulation and cost estimate for biodiesel production using castor oil. Chemical Engineering Research and Design, v. 88, n. 5–6, p. 626–632, 2010.

SOARES, R. DE P.; SECCHI, A. R. EMSO: A new environment for modelling, simulation and optimisation. Computer Aided Chemical Engineering, v. 14, n. C, p. 947–952, 1 jan. 2003.

SOARES, R. EMSO Manual. v. 0300, p. 135, 2007. Available in: https://www.enq.ufrgs.br/alsoc/download/emso/docs/Manual.pdf. Acessed in: August 2020.