



Intelligent modelling and optimisation strategies for the BTO process (Bioethanol-To-Olefins)

Ph.D. Thesis presented by: Gorka Sorrosal Yarritu

Supervised by:
Ph.D. Ainhoa Alonso Vicario and
Ph.D. Cristina Martin Andonegui

within the doctoral program of Engineering for the Information Society and
Sustainable Development



Intelligent modelling and optimisation strategies for the BTO process (Bioethanol-To-Olefins)

Ph.D. Thesis presented by: Gorka Sorrosal Yarritu

Supervised by:
Ph.D. Ainhoa Alonso Vicario and
Ph.D. Cristina Martin Andonegui

within the doctoral program of Engineering for the Information Society and
Sustainable Development

Ph.D. Student

Supervisors

Bilbao, May 24 of 2018

AGRADECIMIENTOS

Ahora que se acerca el final de este largo camino, no puedo evitar mirar atrás y recordar a toda la gente que me ha estado apoyando e impulsando en todo momento en la buena dirección.

En primer lugar me gustaría agradecer a mis Directoras Ainhoa y Cristina, sin cuya ayuda, mucha paciencia y orientación no habría sido posible completar este trabajo. También a mis otros dos “directores en la sombra”, Cruz y Ana, quienes siempre han estado aportando ideas y soluciones durante toda esta aventura. Y tampoco puedo dejar de agradecer a Eloy, sin cuyo apoyo y consejos tampoco habría llegado a buen puerto.

Gracias a todos los compañeros y amigos que hemos compartido tantos momentos en DeustoTech: Aitor, Oihane, Tony, Unai, Ivan, Lohitzune, María, Iraia, Alex, Josu, Mikel y Ander. Sin esos cafés algo más extensos de lo normal, sin vuestro apoyo y ayuda tampoco habría sido posible.

También me gustaría agradecer a mi familia y amigos, especialmente a mi Ama, y a todos aquellos a los que os he puesto alguna vez el trabajo de Tesis como excusa: gracias por comprenderlo.

A Flor mi compañera en la vida, especialmente, por su constante apoyo y comprensión. Por compartir este viaje conmigo, sin cuya paciencia y ayuda, su cariño y sonrisa en los malos momentos, esto no habría sido posible.

Mila esker denori.

Gorka Sorrosal

Bilbao, Mayo 2018

Este trabajo ha sido posible gracias a la financiación del Gobierno Vasco (proyecto BIOTRANS; PI2013-40) y de la Beca Predoctoral de la Universidad de Deusto (EPIF 2012).

ABSTRACT

Nowadays it is clear that crude oil is an exhaustive raw material and energy feedstock, whose reserves decreases annually. However, it is still one of the main raw material and energy feedstock in the world. This situation makes it necessary to develop and industrialise new transformation processes to replace the crude oil as the main raw material.

An important technological alternative solution is based on the use of biomass as the alternative to fossil sources. These technologies are group in the concept of Biorefinery. In this concept are group all the transformation processes, analogous to the petroleum refinery, which using biomass as feedstock are able to produce final marketable products. An important biorefinery process is the transformation process of Bioethanol To Olefins (BTO process). The BTO process, which is the case of the study of the present research work, is a catalytic transformation carried out over an acid zeolite treated catalyst. The BTO process transform a raw material obtained from sources alternatives to petroleum, into key products for the petrochemical synthesis, biofuels and plastics like the olefins C₃-C₄.

An important step towards the industrialisation of this type of processes is the development of advanced and optimised control strategies. The main operating variables of the BTO process are the temperature, the water content in the feed and the space-time. Moreover, the catalyst deactivation becomes more important the longer is the life of the catalyst. In order to improve the production of olefins and make the process economically profitable it is important to develop optimised operation strategies for the process. Both the production rate and the catalyst deactivation rate will fluctuate according to the operational conditions. And the fact that both objectives, production rate and catalyst lifespan, fluctuate in opposite way

depending on the operational conditions, difficult the optimal control strategy.

This work has centred in the modelling and optimisation of the BTO process with the main goal of maximizing the total production of olefins per space-time while at the same time the catalyst lifespan and therefore each production step is prolonged. Both modelling and optimisation methodologies are based on computational intelligence techniques. And therefore, proposed strategy could be used for any new and unknown chemical processes, properly characterized with a representative set of experimental data, but in early experimental stages, saving costs and time.

Soft-computing modelling techniques such as Artificial Neural Networks and Support Vector Machines have been used to model the BTO process. Two modelling approaches have been used, a global modelling and a hybrid strategy which combines soft-modelling techniques with knowledge-based models. In both cases, experimental data of the BTO process and augmented synthetic data generated with a knowledge model of the process have been used during the modelling procedure. The obtained model has been compared with a well-known mechanistic model of the process, which has been used as a contrast method.

In order to optimise the operational conditions of the process, three main dynamic optimisation strategies are proposed. The first one is the optimisation of constant set-points, which is the first proposed operation policy, is one the most common control policy in the chemical industry. The second proposed scenario is the operation using optimised fixed-shape temperature trajectories. And the third proposed strategy is the use of dynamically generated trajectories of all the operating variables generated by the optimised Artificial Neural Networks. These three operation strategies have been optimised using evolutionary algorithms as optimisation technique. During the optimisation procedures, the catalyst deactivation is dynamically considered, being its lifespan extension one of the optimisation objectives. The obtained results are compared with the

operation of the BTO process with constant set-points optimised at zero time of stream.

Finally, the operational conditions are analysed in the last chapter. Different scenarios depending on the cost of the catalyst regeneration phase are presented. The translation of the most suitable operational condition depending on this cost is studied and presented.

TABLE OF CONTENTS

| | |
|---|-----|
| AGRADECIMIENTOS..... | V |
| ABSTRACT | VII |
| TABLE OF CONTENTS | 11 |
| FIGURES | 14 |
| TABLES | 18 |
| SCIENTIFIC CONTRIBUTIONS..... | 21 |
| INTRODUCTION | 23 |
| 1.1. Context and motivation..... | 26 |
| 1.2. Scope of the research and case of study (BTO process) | 27 |
| 1.3. Memory structure | 32 |
| STATE OF THE ART | 35 |
| 2.1. Kinetic modelling of chemical processes..... | 37 |
| 2.1.1. Advanced modelling techniques | 38 |
| 2.2. Optimisation of chemical processes..... | 44 |
| 2.2.1. Evolutionary optimisation | 45 |
| RESEARCH OBJECTIVES | 49 |
| MATERIALS AND METHODS..... | 51 |
| 4.1. Resources: BTO Case of Study | 51 |
| 4.1.1. Mechanistic Model | 51 |
| 4.1.2. Experimental Data | 55 |
| 4.1.3. Synthetic Data | 60 |
| 4.1.4. Soft-Modelling techniques | 63 |
| 4.2. Soft-Modelling methodology | 70 |
| 4.2.1. Data Division..... | 73 |

| | | |
|--------|--|-----|
| 4.2.2. | Model Design and Selection | 73 |
| 4.2.3. | Model Validation | 81 |
| 4.3. | Optimisation methodology..... | 82 |
| 4.3.1. | Constant optimisation | 85 |
| 4.3.2. | Dynamic optimisation..... | 85 |
| 4.3.3. | Dynamic neural optimisation | 87 |
| 4.3.4. | Optimisation at Zero Time of Stream | 88 |
| 4.3.5. | Evolutionary optimisation | 89 |
| | MODELLING OF THE BTO PROCESS | 95 |
| 5.1. | Modelling results..... | 95 |
| 5.1.1. | Linear Regression..... | 95 |
| 5.1.2. | Support Vector Regression | 98 |
| 5.1.3. | Artificial Neural Networks | 102 |
| 5.1.4. | Hybrid Neural Model | 122 |
| 5.2. | Model Comparison | 137 |
| 5.3. | Model Validation | 146 |
| 5.4. | Conclusions..... | 149 |
| | OPTIMISATION STRATEGIES | 151 |
| 6.1. | Optimisation results | 151 |
| 6.1.1. | Optimisation at zero time of stream | 151 |
| 6.1.2. | Constant and dynamic optimisation..... | 153 |
| 6.1.3. | Dynamic neural optimisation | 161 |
| 6.2. | Comparison and discussion | 170 |
| 6.3. | Study of the operational conditions..... | 172 |
| 6.3.1. | Catalyst lifespan extension..... | 172 |
| 6.3.2. | Effect of the catalyst regeneration cost | 178 |
| 6.4. | Conclusions..... | 184 |
| | CONCLUSIONS AND FUTURE WORK | 186 |

| | | |
|--------|--|-----|
| 7.1. | Conclusions and contributions | 186 |
| 7.1.1. | Modelling of the BTO process | 186 |
| 7.1.2. | Optimisation of the operational conditions | 187 |
| 7.1.3. | Actual contributions | 189 |
| 7.2. | Future work and research lines | 190 |
| | BIBLIOGRAPHY..... | 191 |
| | NOMENCLATURE | 205 |
| | ANEX A: MODELLING DATA | 207 |
| | ANEX B: OPTIMISATION RESULTS..... | 213 |

FIGURES

| | |
|--|----|
| Figure 1. World biofuels production trends (Food and Agricultural Policy Research Institute 2011)..... | 24 |
| Figure 2. Biorefinery concept | 25 |
| Figure 3. Transformation scheme of bioethanol into the product lumps of Ethylene, Olefins, Gasoline and Light paraffins (Gayubo et al. 2011). | 29 |
| Figure 4. Diagram of the structure followed by the present memory. | 34 |
| Figure 5. Modelling steps for a chemical reactor. | 36 |
| Figure 6. Example diagram of a hybrid neural model of a chemical reactor. | 41 |
| Figure 7. Transformation of ethanol into ethene (BTE process) and into olefins (BTO process) and other hydrocarbons on acid catalysts (Alonso 2008). | 52 |
| Figure 8. Examples of the operational conditions of the new generated synthetic experiments..... | 62 |
| Figure 9. Data (X) processing example carried out by the ANN (Brierley 1998) | 64 |
| Figure 10. Commonly used activation functions (Brierley 1998) | 66 |
| Figure 11. Structure of a basic ANN (Brierley 1998)..... | 66 |
| Figure 12. Example of a feedforward ANN structure (Demuth et al. 2010)..... | 67 |
| Figure 13. Example of a NARX ANN structure (Demuth et al. 2010)..... | 68 |
| Figure 14. Graphical explanation of the transformation to a higher dimensional feature space (SVR). | 69 |
| Figure 15. Scheme of the followed modelling methodology. | 72 |
| Figure 16. Scheme of the LOOCV procedure for the selection of model structure and its hyperparameters..... | 74 |
| Figure 17. Diagram of the LOOCV training procedure | 75 |
| Figure 18. FIR structural model scheme | 76 |
| Figure 19. NARX structural model scheme..... | 77 |
| Figure 20. ARMAX structural model scheme..... | 79 |
| Figure 21. SSIF structural model scheme | 79 |
| Figure 22. Model design and selection procedure based on prediction capability | 80 |
| Figure 23. Final validation of the selected models..... | 81 |

| | |
|--|-----|
| Figure 24. Feed-Forward ANN topology for generating the trajectories of each operating variable..... | 88 |
| Figure 25. Classification families of the Evolutionary Algorithms..... | 90 |
| Figure 26. Basic principles of the Evolutionary Algorithms..... | 92 |
| Figure 27. Example of the LR based model performance for a constant temperature experiment | 97 |
| Figure 28. Example of the LR based model performance for a variable temperature experiment | 97 |
| Figure 29. Effect of the c and γ parameters on the SVM based model estimates | 99 |
| Figure 30. Effect of the c and v parameters on the SVM based model estimates | 99 |
| Figure 31. Effect of the γ and v parameters on the SVM based model estimates..... | 100 |
| Figure 32. Example of the SVR based model performance for a constant temperature experiment | 101 |
| Figure 33. Example of the SVR based model performance for a variable temperature experiment | 102 |
| Figure 34. Most promising studied ANN configurations | 103 |
| Figure 35. Transfer function during the training procedure | 105 |
| Figure 36. Neural structure selection with the LOOCV procedure for the BASE type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions..... | 107 |
| Figure 37. Neural structure selection with the LOOCV procedure for the DYN type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions..... | 107 |
| Figure 38. Neural structure selection with the LOOCV procedure for the REC type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions..... | 108 |
| Figure 39. Neural structure selection with the LOOCV procedure for the TEMP type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions..... | 108 |
| Figure 40. Neural model architecture (BASE) pre-selected for the BTO process..... | 115 |
| Figure 41. Regression curve of the model BASE (3-5-1) for the entire dataset | 119 |
| Figure 42. Comparison between the simulation and experimental results of olefins concentration when using constant temperature conditions. The simulation results correspond to a mechanistic model (MECH) and neural networks based model (ANN).... | 120 |
| Figure 43. Comparison between the simulation and experimental results of olefins concentration when using temperature ramps. The simulation results correspond to a mechanistic model (MECH), a neural networks based model (ANN)..... | 121 |
| Figure 44. Conceptual scheme of the hybrid modelling approach with ANN..... | 123 |
| Figure 45. Model structures proposed for the hybrid modelling approach. | 124 |
| Figure 46. Neural structure selection with the LOOCV procedure for the MOD 1 type model with linear and log-sigmoid output transfer functions. | 126 |

| | |
|---|-----|
| Figure 47. Neural structure selection with the LOOCV procedure for the MOD 2 type model with linear and log-sigmoid output transfer functions..... | 127 |
| Figure 48. Regression curve of the model MOD 1 (11-23-1) for the training dataset..... | 130 |
| Figure 49. Regression curve of the model MOD 1 (11-23-1) for the test dataset..... | 131 |
| Figure 50. Regression curve of the model MOD 2 (7-20-1) for the training dataset..... | 131 |
| Figure 51. Regression curve of the model MOD 2 (7-20-1) for the test dataset..... | 132 |
| Figure 52. Regression curve of the model MOD 1 (11-23-1) for the simulated data of the original operational conditions | 133 |
| Figure 53. Regression curve of the model MOD 2 (7-20-1) for the simulated data of the original operational conditions | 134 |
| Figure 54. HNN model estimations for the evolution of the BTO process for an experimental run with complex trajectories for the variables T (top panel) and T and X_W (bottom panel) | 136 |
| Figure 55. Prediction capability of the trained models..... | 138 |
| Figure 56. Comparison of the presented model's performance under new complex and unseen operational conditions..... | 141 |
| Figure 57. Comparison of the presented model's performance under new complex and unseen operational conditions..... | 142 |
| Figure 58. Comparison of the presented model's performance under new operational conditions with different prediction horizons..... | 144 |
| Figure 59. Comparison of the presented model's performance under new operational conditions with different prediction horizons..... | 145 |
| Figure 60. Region of similarity and Confidence interval of the TOST test for Real Measurements (top panel) and for the Test simulated data (bottom panel)..... | 148 |
| Figure 61. Feed-Forward NN topology of the hybrid model MOD 2 for main kinetic reaction of the BTO process. | 150 |
| Figure 62. BTO process performance under the operational conditions obtained when optimising the conversion rate of the olefins at zero time on stream..... | 152 |
| Figure 63. Comparison between the different results for the proposed temperature trajectory functions using the MECH model. | 156 |
| Figure 64. Comparison between the different results for the proposed temperature trajectory functions using the HNN surrogate model. | 156 |
| Figure 65. Results of the Friedman test (p-value) and the Nemenyi post-hoc over the different trajectories (Left panel: MECH model; Right panel: HNN model). Linked are the trajectories that are not statistically different attending to the results of the post-hoc test. | 158 |

| | |
|---|-----|
| Figure 66. BTO process performance under the best solution (f_8) provided by the evolutionary optimisation using temperature trajectories (MECH model). | 160 |
| Figure 67. BTO process performance under the best solution (f_8) provided by the evolutionary optimisation using temperature trajectories (HNN model)..... | 160 |
| Figure 68. Comparison between the different results for the proposed neural structures. | |
| | 163 |
| Figure 69. Results of the Friedman test (p-value) and Nemenyi post-hoc over the solutions generated by ANN and trajectories f_8 and f_9 . Linked trajectories are not statistically different attending to the results of the post-hoc test. | 165 |
| Figure 70. Comparison between the different results obtained with ANN (n_6) and f_8 temperature profiles. | 166 |
| Figure 71. BTO process performance and trajectories of the operational conditions for the best solution provide by the evolutionary optimisation ($n = 6$)..... | 168 |
| Figure 72. BTO process performance and trajectories of the operational conditions for and alternative solution provide by the evolutionary optimisation ($n = 8$). | 169 |
| Figure 73. Comparison of the production results obtained with each optimisation strategy. | |
| | 171 |
| Figure 74. Effect of the water content in the feed in the catalyst lifespan extension. | 173 |
| Figure 75. Optimisation results for each fixed X_W value. The red line represents the relation between X_W and the space-time with the maximum TPOST. | 175 |
| Figure 76. Correlation of WF_{EO}^{-1} values for the best solutions for each X_W values. | 176 |
| Figure 77. TPOST maximum values for each combination of X_W and W/F_{EO} values. | 177 |
| Figure 78. Comparison between the best optimisation results obtained for each X_W value. | |
| | 178 |
| Figure 79. Cost of each regeneration procedure depending on the selected penalization index (P_i). | 179 |
| Figure 80. Average TPOST for several optimisation procedures using the proposed penalization indexes..... | 180 |
| Figure 81. Total number of production-regeneration cycles depending on the penalization index..... | 181 |
| Figure 82. Best operational conditions (X_W and WF_{EO}^{-1}) for each X_W value and penalization index. | 181 |
| Figure 83. Results of the grid search (X_W) for each penalization index. Top: Normalized I_C results distribution for X_W . Bottom: Normalized I_C results distribution for WF_{EO}^{-1} | 183 |

TABLES

| | |
|--|-----|
| Table 1. Kinetic constants and the corresponding activation energies for the main kinetic reaction and deactivation process..... | 53 |
| Table 2. Summary of the experimental data..... | 58 |
| Table 3. Examples of some of the most important tested input-output configurations..... | 78 |
| Table 4. Error analysis of MECH model | 82 |
| Table 5. Main parameters of the EA used to optimise the trajectories..... | 94 |
| Table 6. Estimation average errors with the LOOCV procedure for all real experiments with LR based models..... | 96 |
| Table 7. Estimation mean errors obtained for all real experiments with SVR based model | 100 |
| Table 8. Resume of the best internal topology for each structure and transfer function .. | 106 |
| Table 9. Training algorithms used during the neural training procedures..... | 109 |
| Table 10. Comparison of the LOOCV average results obtained for each training algorithm for BASE type model structure | 111 |
| Table 11. Comparison of the LOOCV average results obtained for each training algorithm for DYN type model structure | 112 |
| Table 12. Comparison of the LOOCV average results obtained for each training algorithm for REC type model structure | 113 |
| Table 13. Comparison of the LOOCV average results obtained for each training algorithm for TEMP type model structure | 114 |
| Table 14. Resume of cross-validation estimation average errors for the main implemented models | 116 |
| Table 15. Estimation mean errors obtained for all experiments with MECH, BASE and DYN models | 117 |
| Table 16. Validation average errors with BASE and DYN models for Train, Test and Total datasets | 118 |
| Table 17. Comparison of the LOOCV average results obtained for the most suitable training algorithms for both MOD 1 and MOD 2 models. | 128 |
| Table 18. Resume of the Cross-Validation results for the selected structures. | 129 |
| Table 19. Goodness-of-fit of the models for the Test dataset. | 132 |

| | |
|---|-----|
| Table 20. Goodness-of-fit of the hybrid models to the original operational conditions. | 133 |
| Table 21. Model comparison for all the available real experimental data. | 135 |
| Table 22. Goodness-of-fit comparison for the best trained solution with each modelling technique for the same unseen operational conditions. | 139 |
| Table 23. Results of the goodness-of-fit analysis of the HNN model..... | 146 |
| Table 24. Comparison of the mean production results for the best operating conditions when repeating several times the evolutionary optimisation using the MECH model. | 154 |
| Table 25. Comparison of the mean production results for the best operating conditions when repeating several times the evolutionary optimisation using the HNN surrogate model..... | 154 |
| Table 26. Results of the Shapiro-Wilk test over the production of olefins for each trajectory. | 157 |
| Table 27. ANN training results for the implemented neural structures. | 162 |
| Table 28. Results of the Shapiro-Wilk test over the production of olefins for the proposed neural structures. | 164 |

SCIENTIFIC CONTRIBUTIONS

G. Sorrosal*, E. Irigoyen, Cruz E. Borges, A. Alonso-Vicario (2013). *Modelado del reactor químico del proceso BTO mediante redes neuronales artificiales.* XXXIV Jornadas de Automática, 4-6 Septiembre, Terrasa (Spain)

D. Orive, G. Sorrosal*, C. E. Borges, C. Martin, A. Alonso-Vicario (2014). *Evolutionary algorithms for hyperparameter tuning on neural networks models.* The 26th European Modeling & Simulation Symposium (EMSS2014). September 10-12, Bordeaux (France). The EMSS proceedings are indexed in SCOPUS. ISBN: 978-88-97999-32-4 6.

G. Sorrosal (2015). *Modelado y Optimización inteligente del proceso BTO (Bioetanol a Olefinas).* XIII Simposio CEA de Ingeniería de Control, March 12-13, Donostia-San Sebastián (Spain).

G. Sorrosal*. C. Martin, C.E. Borges, A.M. Macarulla, A. Alonso-Vicario (2015). *Modelado del proceso de transformación de Bioetanol a Olefinas mediante Redes Neuronales Artificiales.* SECAT'15 Catálisis, confluencia interdisciplinar: modelos, catalizadores y reactores. 13-15 Julio, Barcelona (Spain).

G. Sorrosal*, C. Martin, C.E. Borges, A.M. Macarulla, A. Alonso-Vicario (2015). *An optimisation strategy for the catalytic transformation of bioethanol into olefins using computational intelligence.* 15th Conference on Information Technologies - Applications and Theory (ITAT2015). 17-21 Septiembre, Cingov (Slovakia). The CEUR Workshop Proceedings are indexed in SCOPUS. ISBN: 978-151512065-0.

G. Sorrosal*, C.E. Borges, M. Holeňa, C. Martin, A.M. Macarulla, A. Alonso-Vicario (2016). Evolutionary dynamic optimisation of control trajectories for the catalytic transformation of the Bioethanol-To-Olefins process using Artificial Neural Networks. *Genetic and Evolutionary*

Computation Conference (GECCO-2016). July 20-24, Denver, Colorado, (USA). The GECCO Proceeding are indexed in SCOPUS. ISBN: 978-145034323-7.

G. Sorrosal, E. Irigoyen, C. E. Borges, C. Martin, A. M. Macarulla, A. Alonso-Vicario, *Artificial neural network modelling of the bioethanol-to-olefins process on a HZSM-5 catalyst treated with alkali*. Applied Soft Computing, Volume 58, September 2017, Pages 648-656, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2017.05.006>.

G. Sorrosal, C.E. Borges, A. M. Macarulla, C. Martin, A. Alonso-Vicario, *Dynamic optimisation of the Bioethanol-To-Olefins process using Artificial Intelligence*. **Biomass & Bioenergy**, (Submitted).

INTRODUCTION

Nowadays it is clear that crude oil is an exhaustive raw material and energy feedstock, whose reserves decreases annually. However, it is still one of the main raw material and energy feedstock in the world and will continue being for the following decades (Festel et al. 2014). Several governments are pushing for the production of alternatives raw materials. The European Union has established as one of its objectives for 2020 that the 10% of all petrol and diesel transport fuels will have to be from bioresources (Festel et al. 2014).

In nature, 170 billion tons of biomass are produced per year, of which only 3-4% is used, and a large part is therefore available to be used as raw material for the production of intermediate chemicals (case of BTO process) and biofuels (Huber & Corma 2007). The production of biofuels is increasing overall the world in recent years. Figure 1 shows the world biofuels (Ethanol and Biodiesel) production trend (Solgaard et al. 2011; Food and Agricultural Policy Research Institute 2011). As can be see the production of biofuels has increased in the last years and is expected to continue growing in the future.

However the substitution of petroleum-based hydrocarbons and derived raw materials is still a distant goal. The alternative technology necessary to induce the change, is still under study with recent advances in the optimisation and industrial scalability of biomass transformation processes (Jong & Jungmeier 2015). In this context, the price of alternative non petroleum-based products is a decisive factor for the acceptance of them by the markets. A competitive production costs is imperative to extend the biofuels as an alternative to fossil sources (Festel et al. 2014).

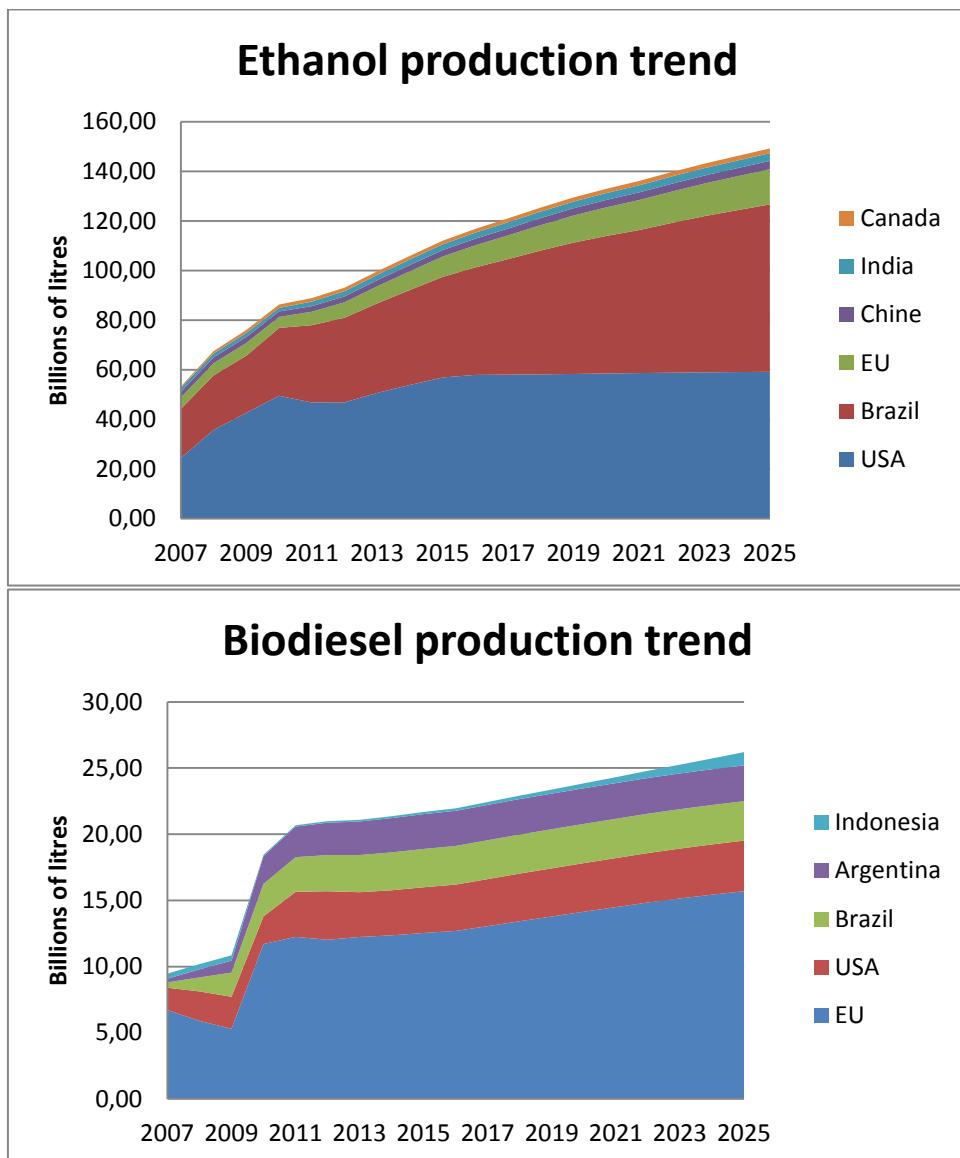


Figure 1. World biofuels production trends (Food and Agricultural Policy Research Institute 2011).

These alternative technologies are group in the concept of Biorefinery. Figure 2 shows the equivalence of the biorefinery with the traditional refinery processes. The biorefinery concept groups all the transformation processes, analogous to the petroleum refinery, able to separate biomass resources as lignocellulosic into their building blocks to be converted into value added final marketable products such as biofuels, energy and/or

other (bio)products and chemicals intermediate products used as raw material for other secondary processes (Koçar & Civa 2013).

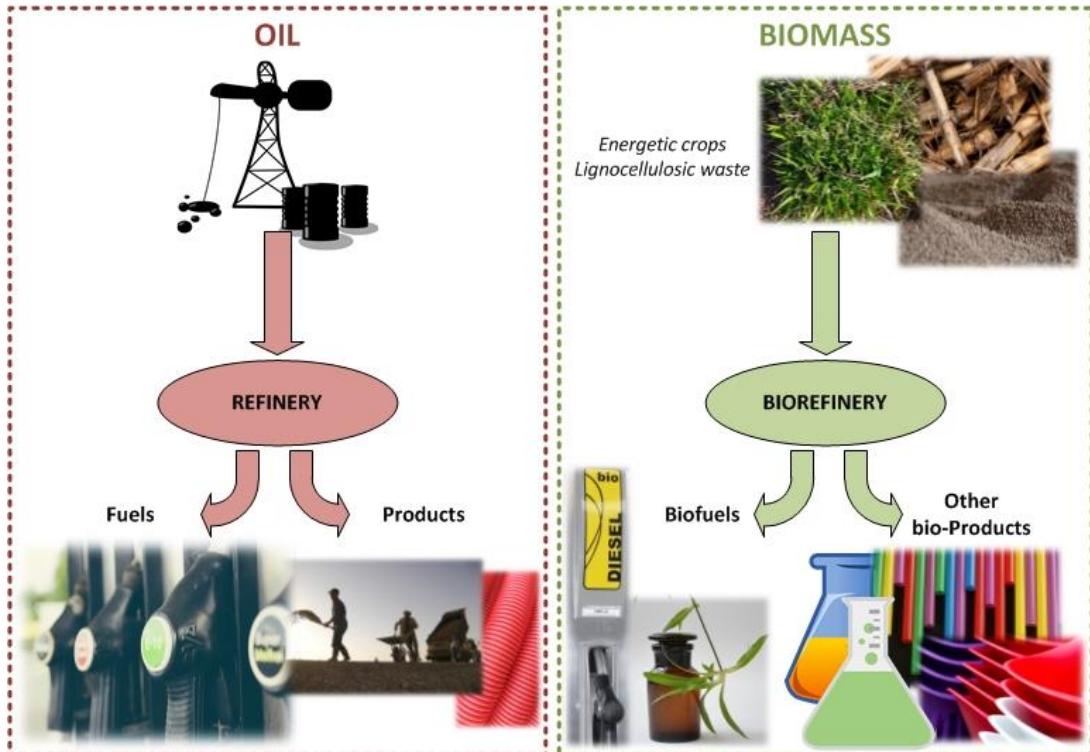


Figure 2. Biorefinery concept

Currently, several countries are promoting the use of biorefinery renewable products. For example, the European Union has set the objective of a 10% of biofuels used for transportation by 2020 and the United States has set a target of 136 billion litres of biofuels by 2022 (Landolina & Maltsoglou 2017). In this context, for transportation, the biofuels is expected to supply up to the 27% of the world fuel demand by 2050 (International Energy Agency 2011).

In order to develop economically viable biorefinery processes, it is necessary to implement new advanced biorefineries with multiproduction purposes and not only focused in biofuels. However, this requires the development of optimal biomass transformation processes in order to minimize the feedstock requirements, maximizing the production per

feedstock unit and strengthening economic viability of the final marketable products. Thus the process automation and expertise in the chemistry, catalysis and engineering areas is necessary for the future implementation of cost-effective hydrocarbon biorefineries (Jong & Jungmeier 2015) or for the adaptation of the existing petrochemical refineries to process biomass derived feedstocks (Huber & Corma 2007).

Although the current relatively low petroleum prices will delay the implementation of advanced new generation biorefinery industries, the expected revenue of US\$10-15 billion for the global chemical industry in the production of biobased products make it possible a future implementation (Jong et al. 2012).

1.1.Context and motivation

In this context the most promising biorefinery processes are not still completely developed and installed (Jong & Jungmeier 2015). Therefore new research works are required to scale-up optimally the biomass transformation processes from laboratory to industrial scale. The development of new tools for a fast advance in the implementation and optimisation of chemical processes, and specially biomass conversion processes, are necessities. It is very important the scientific development in this research area looking for the future post-petroleum society in order to advance in a sustainable society and to reduce the current dependency on the petroleum and its derivate product. Important efforts are being made to reduce the dependence on oil, and increasing the production bioethanol from sustainable sources as the lignocellulosic biomass fermentation (Mosier et al. 2005).

Most recent research works are focused in the design of experiments for a rapid research advances in different chemical processes (Rodemerck et al. 2004), new catalytic material designs (Serra et al. 2007; Valero et al. 2009; Soltanali et al. 2014) or design, control and scheduling of multi-product

chemical processes (Koller & Ricardez-sandoval 2017). These previous works are more centred in the minimization of the experimentation, estimating the relationships between the catalyst composition and their catalytic performance for the process production objectives; or in the whole catalyst performance optimisation to maximise the desired production objective. However, do not consider the dynamic optimisation of the operational conditions of chemical processes. Only Koller and Ricardez-Sandoval (2017) explore a semi-dynamic optimisation of constant step set-points for each sub-product production region.

However, the dynamic optimisation of the operational conditions in complex chemical processes is a different research topic with important improvement possibilities. In complex chemical processes with different kinetic reactions taking place at the same time, the selection of the optimal operational conditions is critical for both maximising the total production and also to select the most suitable catalyst for each process.

Biorefinery processes are complex processes still under study which future implementation depends on their optimisation to obtain final marketable products with reasonable prices. Therefore, the biorefinery future implementation and scalability depends on the research and development of optimised biomass transformation processes.

1.2. Scope of the research and case of study (BTO process)

An important biorefinery process is the transformation process of Bioethanol To Olefins (BTO process). This is a key process in the concept of sustainable refinery, incorporating biomass (or derivatives) as an alternative feedstock to petroleum. The Bioethanol-To-Olefins (BTO) process, which is the case of the study of the present research work, is a catalytic transformation carried out over an acid zeolite treated catalyst (HZSM-5 zeolite treated with alkali NaOH diluted solution (Gayubo, Alonso, Valle, Andrés. T. Aguayo, et al. 2010)) in order to obtain high yields and selectivity

of hydrocarbons. One of the main objectives of the catalyst design was to maximise the production lump of the olefins (propene and butenes). The BTO process transform a raw material obtained from sources alternatives to petroleum, into key products for the petrochemical synthesis, biofuels and plastics like the olefins C₃-C₄.

The olefins (alkenes) are rarely produced in the nature, normally are produced in the petrol processing at the cracking process or from a catalytic process of alkanes (Dossumov et al. 2012). The growing demand of olefins is facing difficulties with the petrol feedstock and with technological limitations of the thermal steam cracking units (main process for the obtainment of olefins). An optimal BTO process has good short-term viability perspectives, as the MTO process (transformation of methanol into olefins) which is rapidly being installed to obtain olefins from natural gas. The production of olefins has a great interest for the industry related to several petrochemical processes with the objective of maximising the production of propylene (Alonso 2008).

Obtaining olefins from biomass is included in the sustainable refinery concept due to the biomass is a renewable resource, easily exploitable, and with a cero effect on the global warming process. The CO₂ emitted during its combustion is near equal to the quantity that the biomass initially eliminates from the atmosphere, which suppose as previously introduced, a zero effect on the global warming. This point is increasingly important and in the focus of the environmental policy of most countries of the world in order to reduce the “greenhouse effect”.

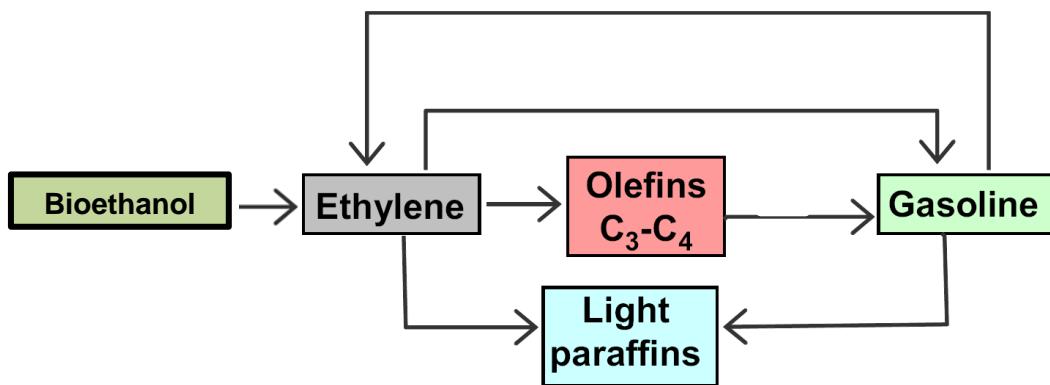


Figure 3. Transformation scheme of bioethanol into the product lumps of Ethylene, Olefins, Gasoline and Light paraffins (Gayubo et al. 2011).

Figure 3 shows the transformation scheme followed in the BTO process. The bioethanol, obtained from the biomass is near completely transformed into ethylene in a dehydration process. Next, the catalyst provides the necessary activation energy to transform the ethylene in three main groups of intermediate products. The products are grouped in four product lumps: the ethylene which is the main process reactant, the lump of olefins which as previously stated is the desired product, and the lumps of gasoline and light paraffins. The objective of the BTO process is to maximise the production of olefins. The total production depends on the used catalyst, which were designed to stimulate the production of olefins and on the operational conditions of the reactor in which the reaction is being carried out. Depending on the operational conditions the correlation between the produced product lumps will vary. The final production will depend on variables as the reaction temperature, the water content or the feeding composition (Bi et al. 2011).

Therefore, it is important to take into account that depending on the selected catalyst, the operational conditions that will maximise the desired production goal will be different (Inaba et al. 2007; Sousa et al. 2016).

The BTO process, as many processes in the current chemical and petrochemical industry and many new processes within the sustainable biorefinery, are carried out in fed bed catalytic reactors. Depending on the

process objective, the reactors are charged with selective and highly active catalysts. However, during the reaction, some catalyst are disable due to their poisoning (or other causes) in a linear or exponential way, or following an unknown function within the deactivation mechanism.

One of the key points for the implementation of this type of process is to perform an advanced process control by adjusting the temperature or other operating variables to counteract the effect of catalyst deactivation extending its lifespan. Due to the fact that the optimal operating points to counteract the deactivation of the catalyst in a commercial plant are limited, because its dependence of multiple variables, it is necessary to develop advanced optimization strategies of the operational conditions that guarantees specific production objectives without exceeding the operation limits to avoid an irreversible deactivation of the catalyst. Therefore, it is very important to have models of the process that allow carrying out simulations or its uses within the controller itself.

Currently, the operation of the BTO process has no an intelligent nor optimised control to improve and maximise the production of olefins. The common operation is limited to establish constant operational conditions, that will not be changed, and proceed with the chemical reaction. Therefore, for the BTO and similar processes as the DTO process (Dimethyl ether To Olefins), the optimisation procedures performed are restricted to the optimisation of the operating variables at zero reaction time (Oar-Arteta et al. 2015; Pérez-Uriarte, Ateka, Gamero, et al. 2016; Sousa et al. 2016) and/or constant set-points (Oar-Arteta et al. 2015; Pérez-Uriarte, Ateka, Gamero, et al. 2016; Da Ros et al. 2017). If the catalyst is deactivated due to the accumulation of coke or other causes, the production phase is stopped and a catalyst regeneration phase is performed, after which a new production phase started. Therefore the objective is to promote the catalyst selectivity of the desired product lump at zero time. This strategy is more useful with chemical processes without catalyst or when the catalyst stability is so high that the deactivation kinetic does not affect to the production in the short-time.

However, in the present case of study (BTO process), depending on the operational conditions the catalyst deactivation rate varies significantly (Alonso 2008). And the operational conditions needed to maximise the production of olefins at zero time are the same which provoke the fastest catalyst deactivation (Alonso 2008; Gayubo et al. 2012). Therefore, with a dynamic optimisation of the operational conditions, the desired production could be significantly improved, looking for equilibrium between maximising the production without inducing a rapid catalyst deactivation or even an irreversible catalyst deactivation.

Given that the optimal operating set-points to counteract the catalyst deactivation are limited in an industrial chemical plant, due to the dependency on multiple variables, it is necessary to develop an operation optimisation strategy so flexible enough to guarantee different production objectives specified by the producers or clients. The relation between the operating variables and the catalyst deactivation can hinder the development and implementation of traditional strategies. Thus, it is necessary the search and development of new strategies for both optimisation and control of biorefinery processes using computational intelligence based techniques, in order to maximise and prolong the production steps, minimising the catalyst regeneration stops and prolonging the total catalyst lifespan.

As previously mentioned, an important inconvenient in the search of the optimal operational conditions is the fact that the kinetics of the main chemical reaction and the kinetic of the catalyst deactivation (to prolong its lifespan) works in an inverse way one from the other. Therefore, it is necessary to find an equilibrium between the process set-points that maximise the production of olefins and the process operation that rapidly or irreversibly deactivate the catalyst.

1.3. Memory structure

The proposed methodology for the optimisation of the operational conditions of the BTO process can be divided in to separate tasks: the modelling and the optimisation procedure itself. Thus, this memory is clearly divided in two main works. First the modelling procedures with the followed methodology for an advanced modelling of the BTO process are presented. Secondly, the optimisation methodologies tested in order to achieve the optimisation goals and the integration of the previous developed model in the optimisation procedure is presented.

Figure 4 shows this division along all the memory structure in order to clearly explain the order and the research carried out in this work. As it can be observes in the diagram, the work and the memory structure is initially divided between the modelling and optimisation procedure and joined for the final optimisation where the model previously developed is used within the optimisation procedure.

Chapter 1 exposes the context and motivation of the work. It presents the BTO process and its importance for a future sustainable society, especially within the current fight against the global warming. In this context, it is exposed the optimisation necessities of the biorefinery processes in order to do it industrialisable and economically viable.

Next, in the Second Chapter, the current state of the art for the modelling and optimisation of chemical processes, giving especial attention to advanced modelling and optimisation techniques, is presented.

Based on the current state of the art, in the Chapter 3 the research objectives and the initial research hypothesis are presented. From this hypothesis the concrete objectives which have guided this work were extracted.

Chapter 4 presents the material and methods used along the work. The experimental setup with the experimental data used along the research work to model and optimise the BTO process is exposed. Moreover, in this chapter a well-known mechanistic kinetic model of the BTO process is presented and the additional synthetic data generated with it. This model will be used during the experimentation for control purposes. Finally, the main principles of the used modelling and optimisation techniques and the both methodologies followed for the modelling and optimisation procedures of the chemical process are presented.

Chapters 5 and 6 exposed the obtained results during the modelling and optimisation procedures respectively. The results obtained with each modelling technique and optimisation strategy presented in Chapter 4 are exposed.

Finally, Chapter 7 presents the conclusions obtained from the previously presented results for both modelling and optimisation procedures applied to the BTO process. Moreover, from the obtained results additional conclusions are extracted for the operation of the BTO process looking for a future industrial scale-up. Finally the future work and possible new research lines are presented.

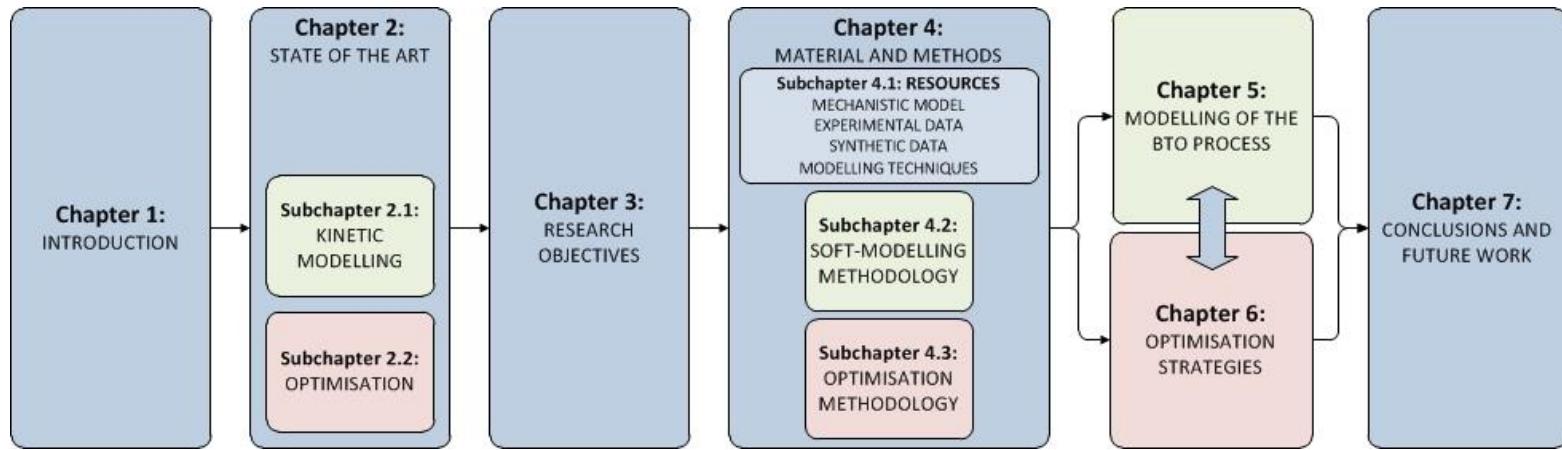


Figure 4. Diagram of the structure followed by the present memory.

STATE OF THE ART

Having reliable and accurate models of chemical processes in general, and for the BTO process particularly, is very important to design, optimises and/or simulate control systems of them. Depending on the needed features such as the accurate or the computational cost of the models, it could be integrated in the controller itself to simulate and estimate the behaviour of the process.

Mathematical modelling of a chemical reactor consist on obtaining the set of equation that represent the physically and chemically the system. Depending on the methodology employed to build it, two types of mathematical models can be distinguished:

- Empirical models
- Mechanistic models

The first one it is based on the results obtained from several test carried out within a specific applicability range of the desired model. These models relate the behaviour of the respond of the systems for the corresponding inputs, without taking into account the reason or explanation of that behaviour. On the other hand, the mechanistic models, more difficult to develop, imply to model the chemical and physical processes that take place inside the reactor. For which it is necessary an expert knowledge of the desired process. This type of modelling implies high mathematical complexity and a huge number of parameters. Thus, mechanistic models usually include some simplifications why they are called semi-empirical models. Figure 5 shows this fact, where the arrows means jumps in the complexity of the model description. Therefore, the final model can be more or less “mechanistic”.

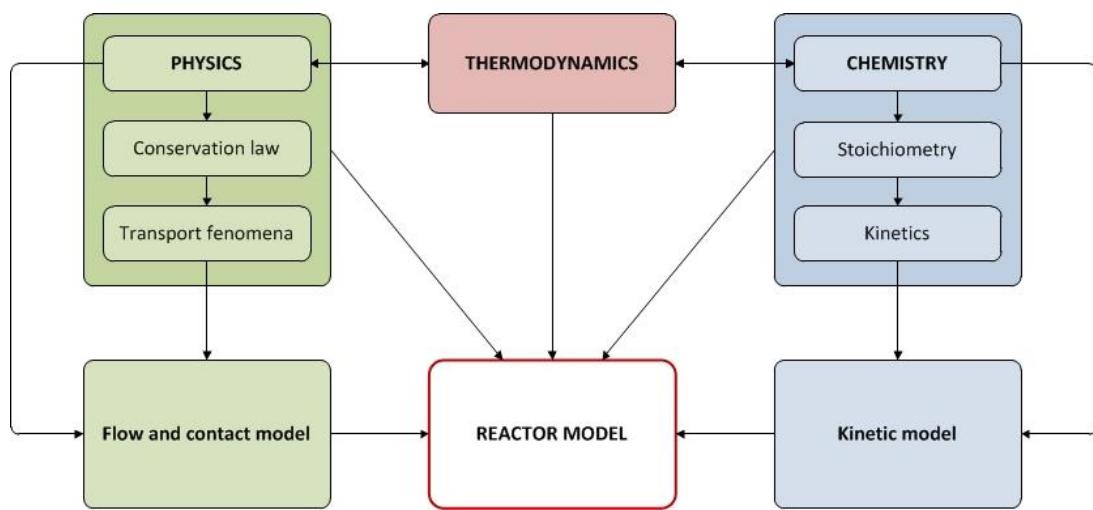


Figure 5. Modelling steps for a chemical reactor.

A mathematical model that completely represents the behaviour of the process should have the following equations:

- Physical and/or chemical equilibrium equations
- Equations of state
- Energy balances
- Mass balance or design equations

However, as previously mentioned, the computational cost of a complete model of the process can be too much for they use for optimisation or control purposes. The modelling and optimisation of chemical processes as the BTO process, has been carried out in previous scientific works. But simultaneous modelling of the kinetics of the main chemical reaction and the catalyst deactivation kinetics is still one of the main problems to obtain a computationally less costly model of the BTO process.

Next subchapters present the most recent advances in the modelling (subchapter 2.1) and optimisation of chemical processes (subchapter 2.2),

focusing on advance modelling and optimisation techniques used with complex chemical processes.

2.1. Kinetic modelling of chemical processes

The development of computational sciences, especially in the last few years, has caused a revolution in the context of the engineering of chemical reaction, permitting a notable decrease in the experimental costs for the development of new processes and products, such as the case of the BTO process. The increase of the scale up to an industrial level, the simulation of the reactor and the optimization of the process, require the knowledge of a rigorous kinetic model. However, the complexity of the mechanisms of catalytic reactions along with the importance of physical transport phenomena, combined with the limitations of the experimental techniques as well as computational methods for the management and analysis of these data, greatly hamper this work (Happel et al. 1990; Atias & de Lasa 2004; Pekar 2004).

Several works are centred in the kinetic modelling, being this one the main difficulties to obtain the desired complete global model of the process. These works are based in established on the one hand on methodologies to develop comprehensive models, introducing a complete chemical and physical phenomena description (Marin & Yablonsky 2011). And on the other hand they use different simplifications based in lumps based reaction scheme proposals to obtain models with the enough compromise between the physical meaning and the accurate significant (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2011; Gayubo et al. 2012; Pérez-Uriarte, Ateka, Aguayo, et al. 2016; Pérez-Uriarte et al. 2017; Toch et al. 2015).

However, the mathematical models of chemical processes are usually complex in their resolution so, even the simplified semi-empirical models, generate very high response times to be used in approaches based on Intelligent Control techniques (Jin 2011). Therefore, a new strategy to obtain

a process model that allows a reduction in the times for the design of the operation strategies and/or the actual control systems is needed.

Due to the complexity of some chemical processes, a new alternative quite extended is to obtain the desired models using artificial intelligence based techniques in the modelling procedures (Cunill et al. 2010). Several techniques such as Artificial Neural Networks are increasingly being used or even linear regression (LR) based methods for simple processes. Cunill (2010) states that the use of these types of techniques is being very effective, it is even the only possible modelling methodology if we are working with complex processes, multiphasic and/or with complex kinetic reactions or unknown reactions (such as the BTO process). The use of these techniques the computational cost of the obtained models is reduced significantly, making it possible their use in optimisation procedures or directly in the process control systems.

2.1.1. Advanced modelling techniques

Several computational intelligence based techniques have been also used to identify and model complex chemical processes (Kajero et al. 2017). The most interesting ones are:

- Artificial Neural Networks (ANN)
- Support Vector Machines (SVM)

The ANN are used in multitude application with successful results due to their learning ability to assimilate high non-linear and time dependent behaviours (Al-Asheh et al. 2007). In addition to their learning capability, it is important to take into account their less computational cost when, once trained, the neural models estimate the behaviour of the process. Their execution time is usually significantly lower than those of the mathematical models of chemical processes. Another advantage of neural modelling is that simplify the modelling procedures of the processes, reducing the

research and developing times and exhibit better performance than the conventional control and modelling techniques (Norgaard et al. 2003).

Last decades, ANN have been widely used for modelling complex systems. As “black-box” models that they are, it is no necessary to have a deep and expert knowledge of the process to be modelled (see subchapter 4.1.4.1). ANN learn from input-output type databases the relation between the input variables and the system outputs. They are consider as universal approximators, being tolerant to fails and noises that appear in the training datasets (Hornik et al. 1989). However, it is necessary to keep in mind that this and the other previously mentioned advanced modelling techniques follows the rule of “garbage in-garbage out”, inherent to all black-box type modelling techniques. This expression wants to denote that the quality of the training datasets employed to implement the models directly and irredeemably affect to the quality and accurate of the developed model (Papadokonstantakis et al. 2005).

In most of the analysed research studies, three layer structure ANN have been used: the input layer, the output layer and only one hidden layer. It has been demonstrated that one only hidden layer long enough can approximate any continue function of n real variables with the desired accuracy (Cybenko 1989; Hornik et al. 1990).

As previously mentioned, due to their characteristics, ANN provide good results when modelling complex processes. Ghadirian and Zekri (2011) tested the performance of neural networks with both single input - single output (SISO) and multiple input - multiple output (MIMO) nonlinear dynamic systems with successful results. Using the resulting model for simulation tasks with the desired accuracy and estimation times (Ghadirian & Zekri 2011). The ability of ANN for modelling, identification and control systems has been widely demonstrated throughout the literature. The ANN, with their different configurations, are capable of modelling all kinds of systems from static systems of a single variable to multivariable systems (MIMO) or dynamical systems, both linear and nonlinear systems (Jagannathant & Lewis 1996; Pollard et al. 1992; Narendra & Parthasarathy

1990). These techniques are already present in industrial sectors such as supply chains, manufacturing industries and several practical applications such as sales forecasting techniques for different objectives, vehicle routing, materials planning or manufacturing flow management (Ko et al. 2010).

In the specific field of chemical processes the ANN have been also used with satisfactory results to predict parameters (soft-sensors) of the production processes as the pH in a continuous flow stirred-tank reactor or to estimate the viscosity of the product; to model steady state reactors; or for the interpretation of biosensor data (Bhat et al. 1990; Bhat & McAvoy 1990; Gonzaga et al. 2009). This first research works, used a global modelling approach where the whole process or reactor is modelled by a unique neural network. However, analysing the literature it has been observed that many of the neural models successfully used in the modelling of chemical reactors are hybrid neural models “Neural Networks-Knowledge Models”, also known as “grey-box” type models, which in contrast with “black-box” models maintain some fundamental knowledge about the process that is being modelled.

These models enable a later simpler reactor scalability (Tian et al. 2001; Hosen et al. 2011). This type of models separate and models on the one hand the most unknown or complex kinetic or thermodynamic reactions of the chemical processes, while maintain in the hybrid model the well-known part of the mathematical model which is easier to resolve and with much less computationally costly (see Figure 6) (Galvan et al. 1996; Molga 2003; Azarpour et al. 2015).

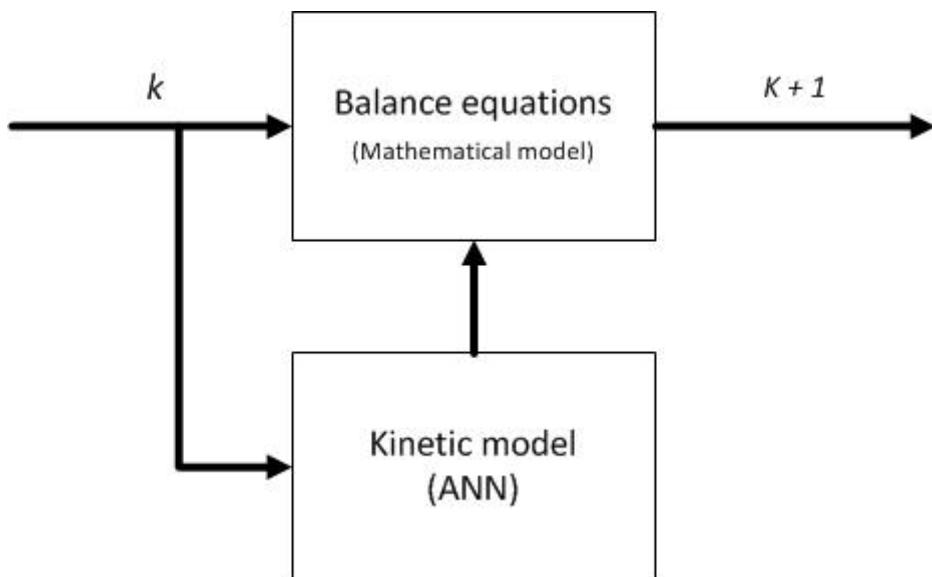


Figure 6. Example diagram of a hybrid neural model of a chemical reactor.

Nevertheless, in the literature, several authors have addressed the problems of chemical process modelling using global neural networks with satisfactory results for different targets. Several complex chemical processes and reactors have been modelled with simulation purposes, such as a catalytic hydrogenation of 2,4-dinitrotoluene performed in multiphase stirred tank reactor (Molga 2003), an industrial fixed-bed catalytic ethylene oxide reactor (Rahimpour et al. 2011) or expanded granular sludge bed biochemical reactors (Yi-Fan et al. 2017). Also less complex but more commonly used batch type reactors have been modelled, or as previously exposed complex kinetic models can be directly modelled (Hough et al. 2017). It is being used the leave-one-out cross-validation approach to obtain the optimal neural structure in order to estimate the output concentrations from several input parameters and previous outputs of the process (Kashani & Shahhosseini 2010). These implemented neural models are able to present better results than well-known mechanistic models, and they are successfully used in different nonlinear control and optimisation strategies (Mujtaba et al. 2006). With a similar approach, a polymerization semi-batch reactor is modelled using a multi-stage modelling approach. Due to a semi-batch reactor works in two different ways, one as a purely batch reaction

and a second one as a semi-batch reaction, with the multi-stage approach several neural models are developed for each process dynamic behaviour. The global model is a stacked neural network in which each neural model, developed individually, assimilates the dynamic of each stage of the process (Yang et al. 1999). With this approach, chemical processes with differentiated dynamic behaviours can be successfully and more easily modelled. Other authors have model polymerization reactors in a global way with different ANN structures, such as the recurrent real-time learning networks or the more common multilayer recurrent networks, and obtaining also successful results (Meert & Rijckaert 1998).

ANN have been also used to model, not only one reactor but also ring reactors or networks of chemical reactors. The methanol production of a ring reactor is predicted by neural models which are able to assimilate the complex dynamics of a group of reactors (Fissore et al. 2004). Other modelling approaches looking for a future control application are present in the literature. An interesting one is the use of ANN to identify the state space non-linear model of an industrial plant (Zamarreño & Vega 1998). The obtained models can be easily linearized, facilitating the application of the conventional control theory.

The previously exposed models of different types are successfully used for control purposes, an important application is the Model Predictive Control (MPC) strategy (Yu & Gomm 2003; Lightbody et al. 1994; Fissore et al. 2004; Sadeghassadi et al. 2018) or fuzzy controllers (Marulanda et al. 2007). In both cases the neural models are used to predict the future behaviour of the process either in an offline mode or even in an online simulation approach.

It is clear that the ANN have been successfully employed for the modelling and control of different several types of reactors and chemical processes. In the reviewed works the authors try always to model the complete behaviour of a specific reactor or process, looking for the relationships between the inputs or the principal components of the inputs

parameters and the outputs of process, generally the output yield of the process or the selectivity of the chemical reaction with the desired product. In all of them very good results have been obtained, using the models in the optimisation and control of the objective process or chemical reactor and/or plants. Nevertheless, ANN have been also used with different purposes within the chemical engineering field. Some of the most important applications found are the experiment design (Serra et al. 2007; Serna et al. 2008) in the research of new materials to reduce experimentation time and efforts (cost) or the modelling of catalytic materials looking for the optimal catalyst composition (Rodemerck et al. 2004; Corma et al. 2005). These works use the neural networks to estimate the catalyst composition and/or material properties or applied treatments to possible catalyst in order to obtain the best performance of the process or catalyst.

The results obtained with ANN are clearly successful. However, they are not the only soft computing technique used for the modelling of commercial, industrial or chemical processes. Other techniques such as Gaussian Processes and SVM have been also successfully used in the literature.

A specific implementation of the SVM, known as Support Vector Regressors (SVR) has been used to model complex systems. The SVR are based in SVM algorithms with some modifications and extensions in order to be used as a regression technique (Smola & Schölkopf 2003; Basak et al. 2007). Different algorithms based on the SVM are used in several applications such as financial markets predictions (Yang et al. 2002), in wind speed prediction for wind farms (Salcedo-Sanz et al. 2011; Santamaría-Bonfil et al. 2016; Hu et al. 2016) or in the electrical industry in which this technique is widely used for short-term, annual and long-term load forecasting (Kavousi-Fard et al. 2014; Wang et al. 2012; Hong 2011).

In the field of chemical engineering, SVM based techniques have obtained in some applications similar accuracy results to those obtained using ANN based approaches (Balabin & Lomakina 2011). They have been successfully used in batch chemical processes using SVM, to build local models around each identified local domain and obtaining a global multi-

model capable to predict the output product concentration (Jin et al. 2015). Good results have been also obtained to model different chemical parameters within chemical processes as adsorption influential parameters or the product quality, and using the estimated information for process monitoring, optimization and quality control (Lee et al. 2005; Ghaedi et al. 2014; Chitralekha & Shah 2010). The modelling of the whole chemical reactors or industrial plants have been also carried using SVR models to predict the process output variables (Bansal et al. 2012; Adib et al. 2013) as well as to model the chemical activity of the processes that are being carried out in the chemical reactors (El-Atta & Hassanien 2017). Also in the petrochemical industry has been used these techniques in recent years (Ahmadi & Pournik 2016).

However, the novelty of the BTO process, whose complete behaviour is still currently under investigation, hinders its modelling using soft computing techniques. The resulting model should adequately reproduce the dynamics of the process, which is the result of both the kinetics of the process itself, and the kinetics of deactivation of the catalyst which promotes the reaction. Therefore, it must be taken into account the two behaviours simultaneously.

2.2. Optimisation of chemical processes

The novelty of the BTO process also makes necessary the study of the most suitable operational conditions for it. The second objective of this research work is the optimisation of the operational conditions of the BTO process. The chemical studies of the operational conditions of chemical processes are usually performed at zero time on stream, and thus they do not take into account the effect of the catalyst deactivation. This is due to the fact that in many chemical processes, such as fermentation processes, the deactivation kinetic is nonexistent or negligible.

The dynamic optimisation of chemical processes at design stages (Kazempour et al. 2017) or to respond to disturbances has been previously approached. It have been performed for yeast fermentation processes in order to optimise the temperature of the reactor (Ławryńczuk 2011) or for biomass fermentation processes for a bioethanol production plant (Ochoa et al. 2010). Relevant works have also been carried out in which the deactivation of the catalyst has been taken into account in a discrete pseudo-dynamic optimisation (Kordabadi & Jahanmiri 2005; Kordabadi & Jahanmiri 2007). In these research works, to optimisation is carried out simply for three or six different states of the activity. In a similar way, some works optimise the processes including the effect of the catalyst deactivation but in a discrete way for a few piecewise of constant profiles (Løvik et al. 1998; Løvik et al. 1999).

The difficulty to tackle simultaneously with two different process dynamics hinder the dynamic optimisation of complex processes such as the catalytic processes. Simple processes and processes with slow deactivation rates can be approached in a discrete way (Løvik et al. 1999). However, in the case of the BTO process, the catalyst deactivation can be very fast (few hours or minutes) depending on the operational conditions, and even a complete irreversible deactivation could happen. Previous research works, have studied the behaviour of the similar catalytic transformation processes of biomass into olefins with different constant set-points and the most suitable constant set-points have been outlined (Pérez-Uriarte, Ateka, Gamero, et al. 2016); nevertheless, to the author's knowledge no dynamic optimisation approaches have been applied to the BTO process.

2.2.1. Evolutionary optimisation

In this context, of complex and intertwined kinetic reactions, the search of the parameters and operational conditions that allow to reach specific production objectives give rise to optimisation problems in which the calculation of an analytic solution can present some difficulties with

conventional search techniques (Yüzgeç et al. 2009). These techniques require characteristics of the process or from the optimisation problem, such as gradients, Hessians or linearities, to calculate the next points. On the contrary, stochastic search techniques, as the Evolutionary Algorithms (EA), solve complex optimisation problems by using stochastic rules.

The EA are inspired on the principles of the biological evolution (Goldberd 1989; Holland 1975). Using properties like the parents selection, the crossovers and the mutations, these techniques have obtained very good results solving complex optimisation problems due to their ability to converge to global optima avoiding being trapped in local optima (Mitra et al. 1998).

In the field of chemical engineering, these algorithms have been frequently employed for the design, optimisation and optimal control of chemical reactors and plants (Kordabadi & Jahanmiri 2005; Angira & Babu 2006; Ghahraloud et al. 2017). The use of these optimisation procedures is not limited to the operational strategy and also includes the design of the chemical reactors itself and the development of catalytic materials. The combination of soft-computing techniques, such as the EA and ANN have been successfully used for the design of new optimised catalyst that improve the selectivity and therefore the production of the desired chemical compound (Serra et al. 2007; Valero et al. 2009); and also to design and guide the experimental set-up of new catalyst in order to experimentally explore the most promising areas (Holeňa & Baerns 2003; Rodemerck et al. 2004).

Some examples of the optimisation of the operational and feeding conditions using and comparing these techniques can be found in the optimisation of the feed flow rate in fermentation processes with fed-batch reactors (Yüzgeç et al. 2009; Yüzgeç 2010) or to optimise the operational conditions of industrial scale reactor (Yee et al. 2003) or chemical plants (Rajesh et al. 2001), both looking for pareto-optimal sets of operating conditions. The multi-objective optimisation is present in several of the previous research works. Most of them using several objective functions to

weigh up depending on the global objective to maximise specific productions or the selectivity of concrete product lumps with conflicting effects. Other authors, search for optimal constant set-points for the production region of each product lump (Koller & Ricardez-sandoval 2017).

One of the main problems when performing a dynamic optimisation of a process or reactor, is the amount of time spent in complete a process simulation. Each evaluation of the cost function may require from minutes to hours of calculation time, and when using EA hundreds of evaluations are generally needed (Ong et al. 2004). Knowledge models describing chemical engineering processes use to be non-linear and with very high computational costs, that is why surrogate models are commonly used to simulate the real process during its optimisation (Nascimento et al. 2000; Laguna & Martí 2002; Muñoz López et al. 2018). Artificial Neural Networks (ANN) are becoming very popular as modelling technique for process simulation in evolutionary optimisations due to their flexibility and good performance (Nascimento et al. 2000; Laguna & Martí 2002; Farshad et al. 2011; Gueguim Kana et al. 2012; Shi et al. 2016; Hough et al. 2017). In addition, the EA have been also used in combination with ANN in neuro-evolution approaches to evolve the neural weighs and structures to obtain the best surrogate models in order to be used during the optimisation procedures (Dragoi et al. 2013; Kajero et al. 2017).

Several studies present modelling and optimising studies of chemical reactors and process operation using soft computing techniques (Angira & Babu 2006; Chen et al. 2004; Kordabadi & Jahanmiri 2007; Babu & Angira 2006; Wei et al. 2015; Ghahraloud et al. 2017). Also in industrial scale different soft computing techniques are increasingly being used with very good results (Velez-Langs 2005; Braik et al. 2008; Oduguwa et al. 2005). However, these studies do not consider dynamically the catalyst deactivation kinetic. There are only some approaches to optimise reactor operation at different discrete activity levels or constant set-point optimisation procedures.

RESEARCH OBJECTIVES

The present research work has the objective of developing a modelling and optimisation strategy based on computational intelligence techniques to be applied on the catalytic transformation of bioethanol into olefins (BTO process). The main goal is to maximise the total production of olefins per space-time while at the same time the catalyst lifespan and therefore each production step is prolonged. However, the proposed strategy could be used for any new and unknown chemical processes, properly characterized with a representative set of experimental data, but in early experimental stages, saving costs and time. There the following hypothesis is performed:

HYPOTHESIS:

The online or offline optimisation of the process using soft computing techniques based on surrogate models, will permit to prolong the catalysts lifespan and to maximise the olefins production; reducing the production stops and costs due to the requirements of the catalyst regeneration phases.

The global objective is therefore, twofold:

O1. Soft-Modelling of the BTO process

The first sub-goal is the modelling of the BTO process using computational intelligence based models such as ANN or SVR. The obtained model of the process should show similar goodness of fit than the knowledge based mathematical models but with much less computational cost. The reduction of the estimation times will make possible the use of the model within optimisation and intelligent control strategies.

O2. Design of the optimisation strategy

The second sub-goal is the design of an optimisation strategy based on the computation intelligence based techniques able to perform a dynamic optimisation of the operational conditions of the processes with the objective of maximizing the total production and extending the catalyst lifespan.

The current studies about operation and optimisation of the BTO process are based on the operation of the process at zero time or constant operation based strategies, which do not able to reach to the desired objectives. This fact can be explained due to the effect that has the catalyst deactivation over the total production. The operational conditions that are able to slow down the catalyst deactivation process are inverse to those necessaries to maximise the production at zero time (elevated temperatures).

The proposed automated modelling and optimisation strategy will make easier and faster to study and improve new processes whose chemical reaction kinetics are still under study and those that do not exist a valid knowledge based mathematical model. The research and development in the field of biorefinery processes is especially important. The obtained advances and conclusions are critical to obtain industrial scale and economically viable productions of bioproducts.

MATERIALS AND METHODS

In this chapter the materials and methodologies used to develop the presented research work are presented. First the experimental setup and data used for the modelling procedure are presented. Next the mechanistic model, which is used for control and contrast purposes, is presented and explained. And finally the methodologies and techniques used for the modelling and optimisation procedures are explained. The modelling methodology including the modelling techniques and the statistical analysis for the model selection and validation procedures are presented. In the optimisation procedures, three different optimisation and process operation strategies are presented with the selected optimisation procedure.

Both modelling and optimisation procedures have been implemented using the programming package MATLAB[©] (version 8.0, 2012b, Mathworks Company) and the libSVM library (Chang & Lin 2011) for the alternative Support Vector Regression. Simulation and optimisation have been carried out in a PC with an Intel[®] Core[©] i5-2467M CPU at 1.6 GHz and 4.0 GB of physical memory (RAM).

4.1. Resources: BTO Case of Study

4.1.1. Mechanistic Model

During the work a previously developed and validated mechanistic model of the BTO process will be used. An implementation of this model, which includes both main reaction kinetics and the deactivation kinetic, is used for contrast purpose during the comparative study of the new developed surrogated models.

The kinetic model proposed by (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012) according to the kinetic scheme of Figure 7 has been chosen and it will be considered as the mechanistic model (MECH). In this model four component lumps of commercial interest are considered: ethylene (E), lump of gasoline (G) (butane and C5+), lump of propylene and butenes (O), and lump of light paraffins (P) (C4–). Ethylene is considered the starting reactant given that ethanol dehydration is complete under all the operating conditions studied.

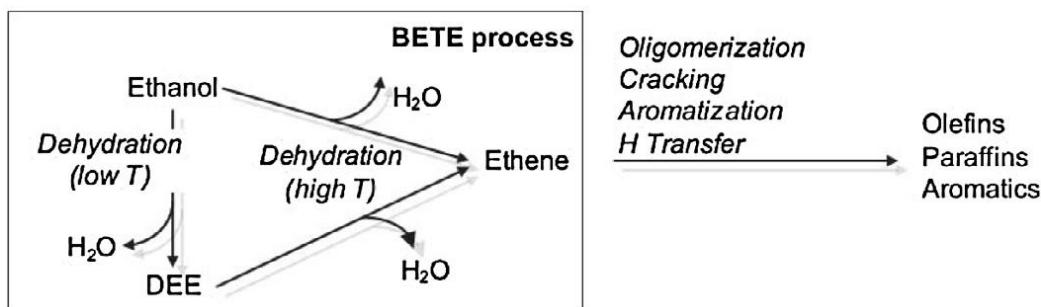


Figure 7. Transformation of ethanol into ethene (BTE process) and into olefins (BTO process) and other hydrocarbons on acid catalysts (Alonso 2008).

The kinetic equations of the main reactions (at zero time on stream) are:

$$(-r_E)_0 = \left(((k_1 + k_2 + (k_2 + k_4)X_O)X_E) - k_6 X_G \right) \cdot \theta \quad (1)$$

$$(r_O)_0 = (k_1 X_E + (k_2 - k_4)X_O X_E - k_3 X_O + k_6 X_G) \cdot \theta \quad (2)$$

$$(r_G)_0 = (k_3 X_O + 2k_4 X_E X_O - 3k_6 X_G) \cdot \theta \quad (3)$$

$$(r_P)_0 = (k_5 X_E + k_6 X_G) \cdot \theta \quad (4)$$

Where r_i is the formation rate of i component at zero time on stream and at any time. The parameters k_i represents the kinetic constants of i step in the kinetic scheme of the model at T temperature. These parameters are expressed using the equation of Arrhenius, which is reparametrized to redefine the kinetic constants to be calculated as:

$$k_i = k_i^* \exp \left[-\frac{E_i}{R} \left(\frac{1}{T} - \frac{1}{T^*} \right) \right] \quad (5)$$

being k_i^* the kinetic constant of i step at the reference temperature T^* of 573 K, which is the lower temperature within the experimental data range. R represents the constant of gases and E_i the activation energy of each i step in the kinetic scheme.

Table 1. Kinetic constants and the corresponding activation energies for the main kinetic reaction and deactivation process.

| Kinetic constants | |
|----------------------------|---|
| k_1^* | $0.626 \pm 0.052 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_2^* | $6.208 \pm 0.517 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_3^* | $2.068 \pm 0.230 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_4^* | $6.042 \pm 0.731 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_5^* | $0.018 \pm 0.003 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_6^* | $0.016 \pm 0.003 \text{ (g/(g}_{\text{catalyst}}\text{h))}$ |
| k_w^* | 0.230 ± 0.007 |
| k_d^* | $0.361 \pm 0.030 \text{ (h}^{-1}\text{)}$ |
| k_{dw}^* | 0.238 ± 0.023 |
| Activation energies | |
| E_1 | $96.85 \pm 3.140 \text{ (kJ/mol)}$ |
| E_2 | $70.60 \pm 3.977 \text{ (kJ/mol)}$ |
| E_3 | $54.43 \pm 4.229 \text{ (kJ/mol)}$ |
| E_4 | $59.81 \pm 4.647 \text{ (kJ/mol)}$ |
| E_5 | $148.30 \pm 7.536 \text{ (kJ/mol)}$ |
| E_6 | $130.40 \pm 6.238 \text{ (kJ/mol)}$ |
| E_w | $24.35 \pm 1.210 \text{ (kJ/mol)}$ |
| E_d | $74.70 \pm 2.17 \text{ (kJ/mol)}$ |
| E_{dw} | $8.941 \pm 0.300 \text{ (kJ/mol)}$ |
| Parameters | |
| n_d | 1.845 ± 0.077 |

Table 1 shows the values of the kinetic constants for a reference temperature (573 K) and the corresponding activation energies, for the main kinetic reactions and the deactivation kinetic by coke deposition. The function $\theta(X_w)$ quantifies the attenuation of individual reaction rates due to the presence of water in the reaction medium according to the expression:

$$\theta = \exp(-k_W X_W) \quad (6)$$

where X_W , is the water content in the reaction medium, which is calculated as the sum of the water fed together with ethanol, X_{W0} , and that formed by ethanol dehydration, X_{Wf} .

The content of formed water is given by:

$$X_{Wf} = \frac{18}{28} \left(\frac{g_{H_2O}/mol_{EtOH}}{g_{C2=}/mol_{EtOH}} \right) = 0.643 \left(\frac{g_{H_2O}}{g_{C2=} \text{ in the feed}} \right) \quad (7)$$

It should be taken into account that the mass of ethylene equivalent in the reactor feed equals the mass of the organic components in the reaction. Furthermore, the mass conservation equation for each i component lump considering plug flow in the fixed bed reactor, expressing the longitudinal coordinate as dimensionless, ζ , is:

$$\frac{(\delta X_i)}{\delta t} = \frac{(1 - \varepsilon)\rho R \cdot T}{\varepsilon \bar{M}} (r_i)_0 \left(\frac{m_0}{F_{EO}} \right) a - \frac{u}{Z} \frac{\delta X_i}{\delta \zeta} \quad (8)$$

where the activity a is defined as the ratio between the reaction rates at t and zero time on stream:

$$a = \frac{r_i}{(r_i)_0} \quad (9)$$

The deactivation kinetic by coke of the catalyst considers that coke formation occurs independently of reaction medium concentration, according to the expression:

$$\frac{da}{dt} = -k_d \theta_d a^{n_d} \quad (10)$$

The term k_d is the kinetic constant for the deactivation by coke formed at temperature T . In the same way as for the kinetic constants of the main kinetic reaction, the deactivation kinetic constants are calculated with the

reparametrized equation of Arrhenius (equation (5)) from the deactivation kinetic constant k_d^* at the reference temperature T^* .

Finally, θ_d quantifies the attenuating effect of the water in the reaction medium on deactivation by coke according to an exponential expression:

$$\theta_d(X_W) = \exp(-k_{dW}X_W) \quad (11)$$

being the term k_{dW} the kinetic constant for the deactivation by coke formed from the water content in the feed.

It is therefore a multivariable process with strict operating restrictions. The possible operational conditions are bounded based on the physical-chemical properties of the process. In fact, for the temperature (T), 573 K is the inferior bound where the complete dehydration of the ethanol happens and 673 K is the upper bound to avoid an irreversible deactivation of the catalyst. The variable X_W will range between 0.0821 and 4.8889 $\text{g}_{\text{water}}\text{g}^{-1}$, and W/F_{EO} will range between 0.068 and 1.525 $\text{g}_{\text{catalyst}}\text{h}(\text{g}_{\text{ethanol}})^{-1}$ respectively. Please note that those intervals have been used in previous works (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012).

4.1.2. Experimental Data

This study analyses the olefin composition (X_O , propylene and butenes) as the main reaction product, taking into account the following main variables and disturbances (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010):

- Operating variables:
 - T: reaction temperature (K).
 - X_W : mass fraction of water based on the equivalent mass of ethylene in the reactor feed ($\text{g}_{\text{water}}\text{g}^{-1}$).
 - W/F_{EO} : space-time ($\text{g}_{\text{catalyst}}\text{h}(\text{g}_{\text{ethanol}})^{-1}$).

- Disturbances:
 - a: catalyst activity. The activity has been considered as a disturbance that quantifies the rate of catalyst reversible deactivation by coke deposition.

The experimental data were obtained, in previous works (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012), in an automated device equipped with an isothermal fixed bed reactor connected on-line to a gas chromatograph (Hewlett Packard 6890 Series II) and a micro-GC (Varian CP) for the analysis of the reaction products. The reactor consists of 316 stainless steel with a 9 mm of internal diameter and 10 cm of total effective length located within a stainless steel chamber heated by electrical resistances. The bed is a mixture of a HZSM-5 catalyst with an alkali treatment of 10 min and inert sand in order to ensure an isothermal bed and the plug flow under low space-time conditions. The experiments have been obtained under the following operating conditions: temperature, between 573 and 673 K; pressure, 1 atm; space-time, up to $1.53 \text{ (g}_{\text{catalyst}}\text{)}\text{h}(\text{g}_{\text{ethanol}})^{-1}$; particle size of the catalyst, between 0.3 and 0.63 mm; feed (ethanol + water) flow rate, $0.16 \text{ cm}^3\text{min}^{-1}$ and $30 \text{ cm}^3\text{min}^{-1}$ of N₂; water content in the feed, up to 75 wt%; and time on stream, up to 31 h.

Table 2 summarizes the experimental dataset used in this work. Details about the reaction equipment, catalyst preparation and experimental methodology can be found in previous works (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012). Two types of experiment are distinguished according to the reaction temperature: experiments at constant temperature (T_{const}) and dynamic experiments with a temperature ramp (T_{ramp}). The purpose was to have the greatest quantity of information about the process dynamics with the minimum number of experiments.

Each of the experimental samples is a vector of data featuring the conversion of olefins at certain operating conditions. The small number of samples with respect to the variety of experimental operating conditions greatly hinders the modelling process because of the dependence of the

modelling techniques (Papadokonstantakis et al. 2005) with the quality of training data.

Table 2. Summary of the experimental data

| | T [K] | T slope [K/min] | X _W [g _{water} g ⁻¹] | W/FEO [g _{catalyst} h(g _{ethanol}) ⁻¹] | Duration [h] | Output samples (X _O) |
|--------------------------|----------|--------------------|---|--|-----------------|-------------------------------------|
| T_{const} | 573 | 0.0821 | 0.0821 | 0.235 | 13.5 | 11 |
| | | | 1.643 | 0.235 | 14.2 | 16 |
| | | | 1.643 | 0.469 | 14.2 | 16 |
| | | | 4.888 | 1.108 | 6 | 18 |
| | 623 | 0.0821 | 0.0821 | 0.132 | 20.45 | 20 |
| | | | 0.0821 | 0.163 | 9.5 | 11 |
| | | 0.0821 | 0.0821 | 0.235 | 12.45 | 14 |
| | | | 1.643 | 0.235 | 15.1 | 16 |
| | | | 1.643 | 0.469 | 15.1 | 16 |
| | | | 4.888 | 0.76 | 30.5 | 27 |
| T_{ramp} | 648 | 0.0821 | 0.0821 | 0.068 | 6.83 | 16 |
| | | | 0.0821 | 0.132 | 22.08 | 16 |
| | | | 1.643 | 0.235 | 15.1 | 16 |
| | 673 | 1.643 | 1.643 | 0.469 | 15.1 | 16 |
| | | | 1.643 | 0.235 | 25.75 | 16 |
| | | | 4.888 | 1.525 | 27.56 | 16 |
| | 573-648 | 0.5 | 0.0821 | 0.132 | 28 | 16 |
| | | | 0.0821 | 0.235 | 3.3 | 11 |
| | 573-673 | 0.5 | 0.0821 | 0.235 | 16.5 | 32 |

| | T [K] | T slope [K/min] | X _W [g _{water} g ⁻¹] | W/F _{EO} [g _{catalyst} h(g _{ethanol}) ⁻¹] | Duration [h] | Output samples (X _O) |
|-----------------------------|----------|--------------------|---|--|-----------------|-------------------------------------|
| 573-673 | | | 1.643 | 0.235 | 6 | 19 |
| | | | 1.643 | 0.469 | 4.33 | 14 |
| | 0.6 | | 1.643 | 0.117 | 2.97 | 16 |
| | | | 0.7 | 0.068 | 7.83 | 19 |
| TOTAL: 23 experiments | | | | | | 388 |

4.1.3. Synthetic Data

The previously presented experimental data will be used to fit different models using soft-computing based techniques. However, during the modelling procedure was noticed that the small number of samples with respect to the variety of experimental operating conditions and the characteristics of the available experimental runs, was not able to provide to the model with the necessary dynamic information about the BTO process. The experimental data is limited to short-time experiments where the catalyst deactivation has a relatively low influence on the final olefins conversion rate. Only 6 experiments, out of 23, represent the evolution of the process more than 16 hours. And in many of the experimental runs (10 of 23 experiments), the catalyst deactivation and therefore its effect over the production is minimal (less than a 25%) due to its short duration.

However, optimizing the operational conditions, several operational conditions are simulated in order to extend the production times and catalyst lifespan, and therefore prolonging the prediction horizon to analyse different scenarios describing different operational conditions. This type of behaviour is not represented in the experimental dataset. Moreover, the dataset does not contain any experiment that considers the dynamic behaviour of X_w and W/F_{EO} and few ones for the T . So it is probably that in order to train an empirical model capable to assimilate the dynamics of the process for long-term estimations more training data may be required.

Therefore, in addition to the experimental data, and in order to provide to the model more dynamic information about the process, 20 additional experiments with different operational conditions have been simulated. These additional data have been generated using the well-known mechanistic model (see subchapter 4.1.1).

These augmented data generated on extended operational conditions will provide the trained models with the necessary dynamic information

about the behaviour of the process in long-term experiments (more than 60 hours) with new dynamic operational conditions of all variables and not only for temperature ramps. In addition, these new data will provide more information about the effect of the catalyst deactivation over the total production of olefins. This effect becomes critical in long-term estimations of the production of olefins.

Figure 8 shows as an example, some of the new generated synthetic operational conditions that will be used to train the final models. The new dataset composed by 43 experimental conditions (the 23 original experimental conditions and 20 additional “dynamic” operational conditions) have been simulated by the mechanistic model and treated hereafter as the synthetically generated experimental data.

Intelligent modelling and optimisation strategies for the BTO process (Bioethanol-To-Olefins)

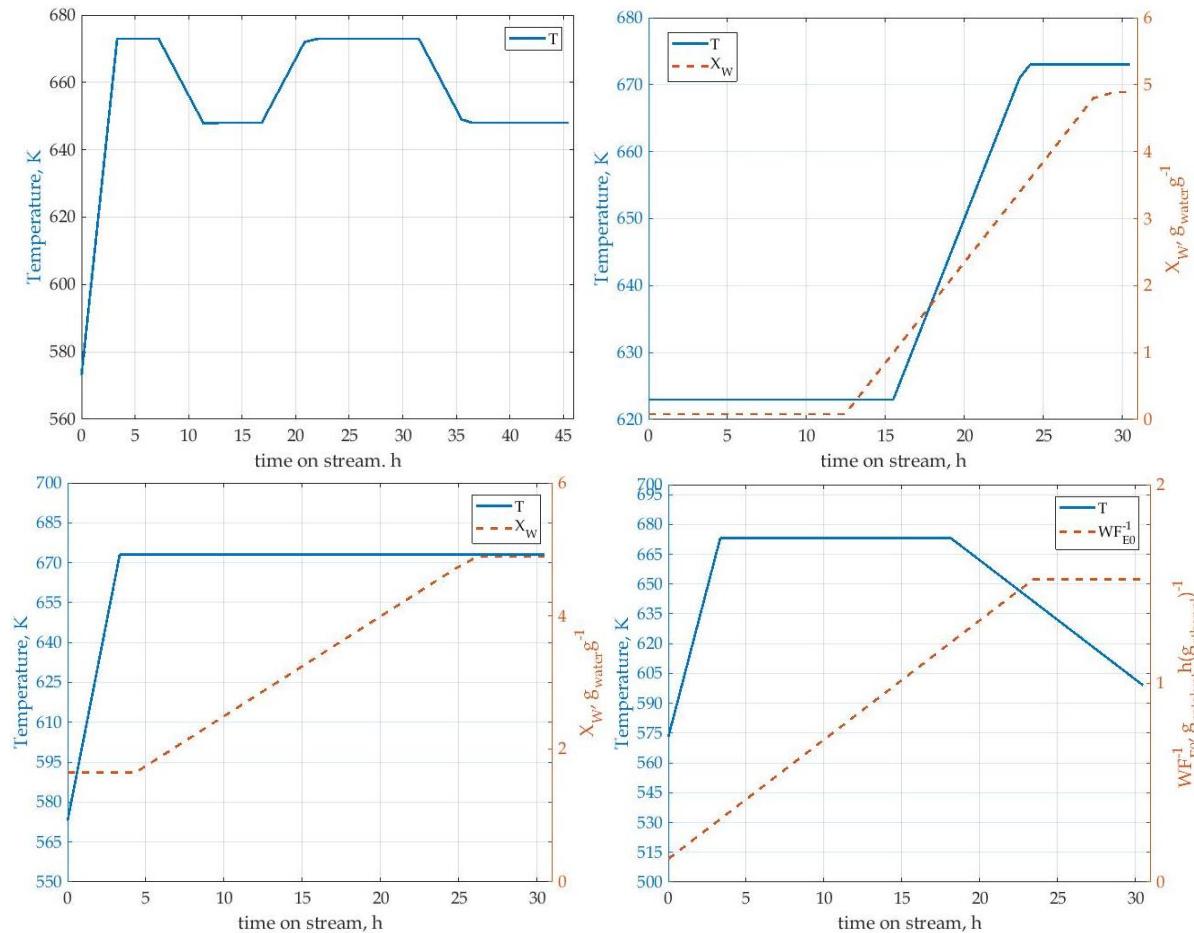


Figure 8. Examples of the operational conditions of the new generated synthetic experiments.

4.1.4. Soft-Modelling techniques

According to the literature one of the most promising soft-modelling techniques is the ANN, and therefore this technique has been selected for the modelling procedure. However, modelling techniques as the Support Vector Machines (SVM) have also been tested due to their promising results in the modelling and identification of complex processes. Additionally, and for contrast purposes, a Linear Regression (LR) technique has been used to fit the process data.

Using the process data previously presented and the commented modelling techniques (ANN, SVR and LR), the modelling of the BTO process has been addressed. However, in addition to the main operating variables of the process, another variable that has a significant effect over the behaviour of the process is the activity of the catalyst. Moreover, the influence of this variable is determinant and increases with the production time. However, the activity is a non-measurable variable and therefore no experimental data is available for the modelling procedure. Therefore, to the previously presented global modelling technique, hybrid modelling approaches will be included in order to be able to assimilate the deactivation and to identify both main kinetic and deactivation kinetic reactions of the BTO process.

Thus, a hybrid modelling approach is also proposed, using soft computing techniques such as ANN to model the main process kinetic and a knowledge model for the catalyst deactivation (see subchapter 5.1.4 for more details).

Next subchapters present each modelling technique used to model the BTO process in this work.

4.1.4.1. Artificial Neural Networks

ANN are commonly used in the modelling of complex systems. This is due to their successful results in the resolution of problems with stochastics and non-linear characteristics, and because of their great learning capacity from previous experiences and generalization ability.

They are classified within the category of “Black-Box” modelling techniques. This means that the model parameters have not a real physical interpretation. The model parameters only correlate the inputs of the model with the corresponding outputs. Therefore it is not necessary a deep knowledge of the process which allows faster progress in some research areas.

ANN use the structural principles of the biological brains to design intelligent systems, and trying to emulate the low level performance of brains. The artificial processing units are called nodes or neurons. These are equivalents to the biological neurons, and try to emulate their behaviour and operation. They transmit the values received in their inputs to their outputs, going from one neuron to another. The data processing that is carried out in each neuron is local and depends on mainly in the data input values and the weights that are associated in each connexion between neurons.

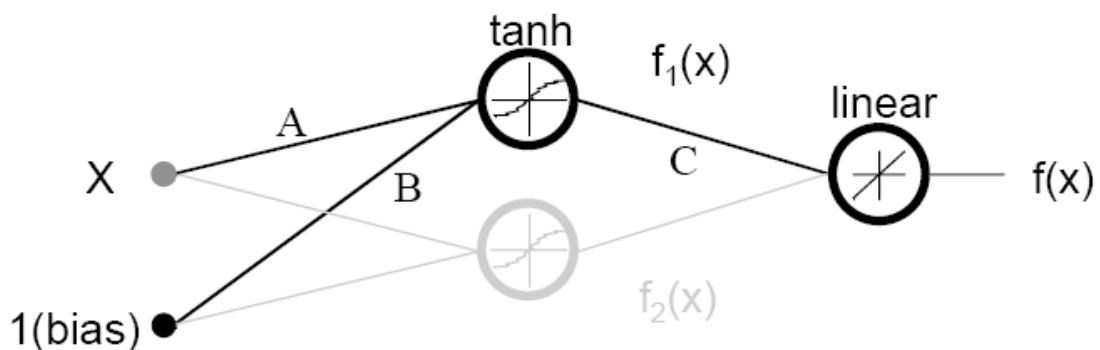


Figure 9. Data (X) processing example carried out by the ANN (Brierley 1998)

Figure 9 shows the operation of neurons to process the data received in their inputs. In this small neural network, the output $f(x)$ will consist of the sum of the entries (which are the outputs of previous neurons), weighted by weights associated to each neural connexion, and applying the activation function corresponding to the current node. Therefore, the output of the ANN will be the next:

$$f(x) = f_1(x) + f_2(x) \quad (12)$$

being,

$$f_1(x) = C \cdot \tanh(Ax + B) \quad (13)$$

The weights associated to each connexion (A, B, C), are used to weight up the inputs of each node of the network. These values will be fitted during the network's training procedure in order to minimize an error function between the estimates generated by the model for an input vector and the objective or real output corresponding to those input vector.

Each neuron has an activation function such as the one that can be seen in equation (13), and which is applied to the weighted addition of inputs. In this example, the activation function is the hyperbolic tangent, but different function types can be used. These functions can be different between neurons and normally the output neurons of the networks used to use linear activation functions. However, the election of these functions depends of the specific use case in each modelling problem. Next, Figure 10 shows the most commonly used activation functions.

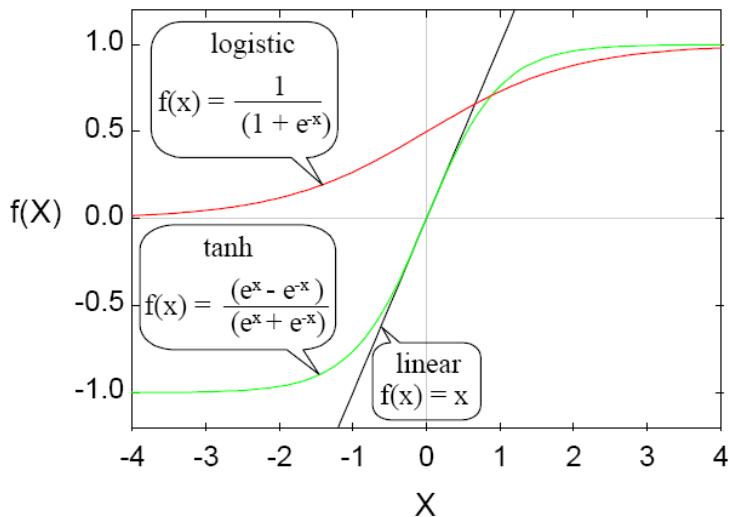


Figure 10. Commonly used activation functions (Brierley 1998)

The structure of the ANN is composed of a variable number of nodes (neurons) organized in different layers that process the inputs in parallel inside each of this layers.

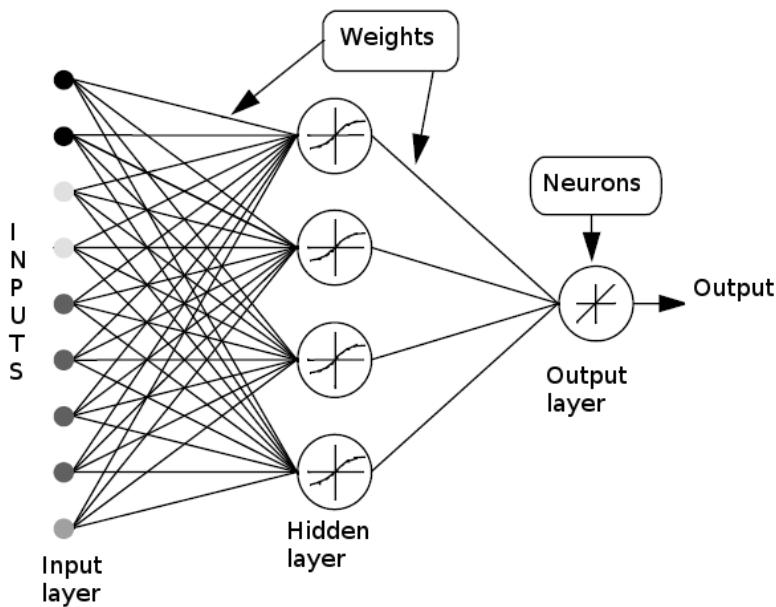


Figure 11. Structure of a basic ANN (Brierley 1998)

Figure 11 shows the three layer types in which can be divided the neural structure:

- *Input layer*: it corresponds with the input to the neural network, and it is where the input data vectors are connected.
- *Hidden layers*: they are the processing core of the neural network. It can be one or several hidden layers with a variable number of neurons in each one.
- *Output layer*: it is the output of the neural network, it corresponds with the outputs of the model.

Depending on the existing connections between these layers or between themselves, neural networks are classified as static or dynamic networks. Static networks (see Figure 12) has no recurrent loops and therefore their outputs are obtained only from the available inputs and following feedforward connections (Demuth et al. 2010).

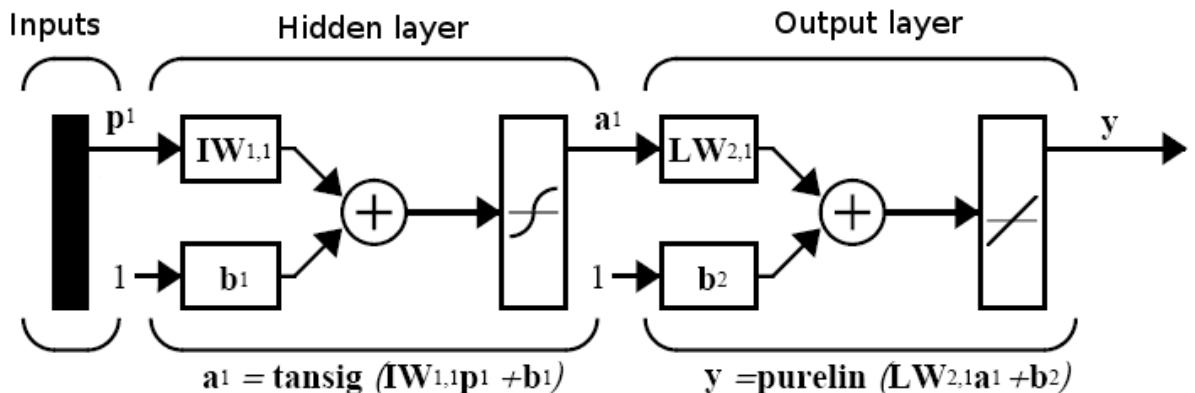


Figure 12. Example of a feedforward ANN structure (Demuth et al. 2010)

In contrast, the dynamic neural networks the outputs are not only estimates from current input data. As can be seen in Figure 13, the output is obtain from current inputs but also from previous neural outputs of the network or other network variables.

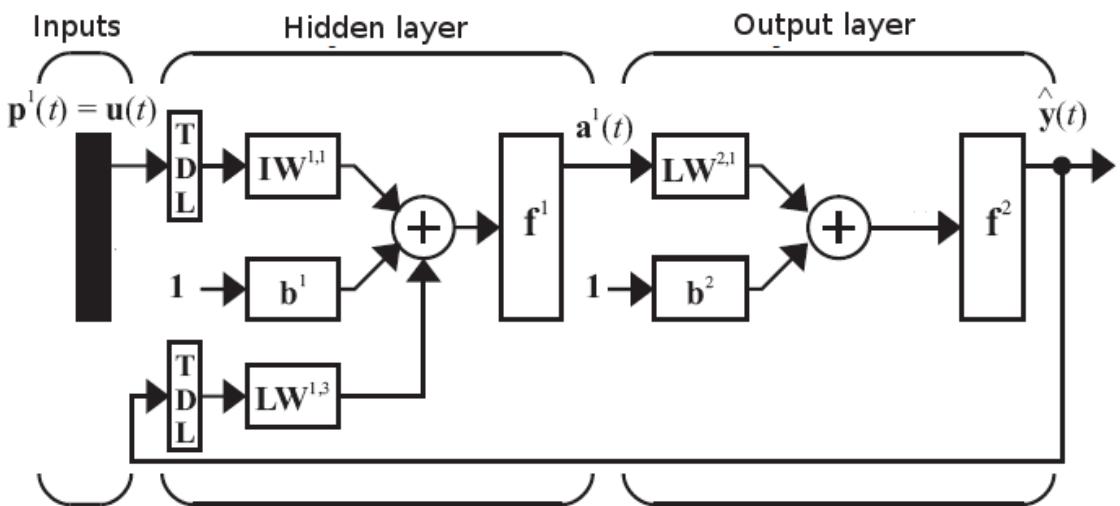


Figure 13. Example of a NARX ANN structure (Demuth et al. 2010)

Finally, a training process is needed in order to fit the weights of the ANN. This point is the most important in the learning capability of the networks. It exist several methods and training algorithms to update iteratively each weight to reduce an error calculated from available output data or from previously seen patterns. All of them can be classified as supervised or non-supervised learning methods.

In the supervised learning methods, the ANN is fed with information about bot inputs and its corresponding outputs. The training algorithm will fit the neural weights with the final objective of minimizing the error between the network output and the real output data. During the training process several correlated input-output data are used to fit the weight until an acceptable error measurement is reached.

In contrast, in the non-supervised learning methods, the ANN receive only information about the inputs. These inputs contain patterns which the network has to learn, recognizing new patterns based on previously processed experiences.

4.1.4.2. Support Vector Machines

The SVM represent another widely known “black box” modelling technique. SVM are a set of supervised learning algorithms originally designed for solving classification problems. With certain modifications, they are being used in recent times with great success for time series regression and estimation applications (Smola & Schölkopf 2003).

The original SVM search for hyperplanes capable of separating the input-output vectors in an optimal form. The objective is to maximize the distance (margin) from the nearest vectors to the hyperplane. In order to achieve this objective, these points are usually projected to a space of superior dimensionality. The vector form by the nearest points to the hyperplane is called support vector.

The Support Vector Regression (SVR) is a specific type of SVM which consist of a non-linear mapping of the input-output training data to a higher dimensional space in which a linear regression could be performed. Figure 14 graphically shows the transformation from a non-linear feature space (X) to a higher dimensional space ($\varphi(X)$) where the modelling problem is transform and now can be solve as a liner regression.

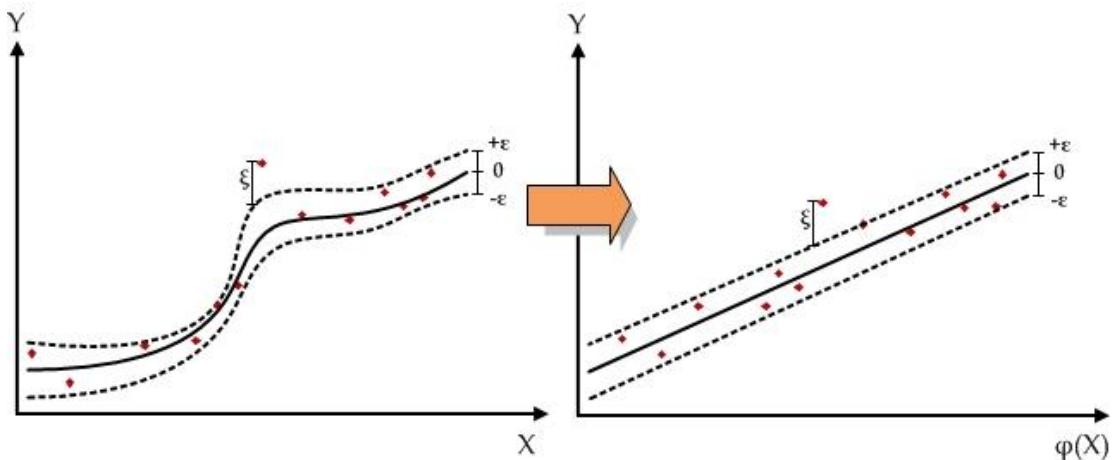


Figure 14. Graphical explanation of the transformation to a higher dimensional feature space (SVR).

4.2. Soft-Modelling methodology

In order to obtain the desired model, an iterative methodology with different modelling techniques has been followed. For each modelling technique several structures and hyperparameters have been tested until the most suitable model has been found and finally validated. Next the modelling methodology used to design, select and train the soft-computing based model is presented.

The same modelling methodology has been applied with each modelling technique. The proposed method consists on an iterative methodology modifying the structure or the hyperparameters of the models. These parameters will be different depending on the selected techniques. In the case of ANN, for each global iteration the number of layers and the neurons in each layer has been updated. For the SVM based models several kernel functions with different hyperparameters for each one have been tested. Figure 15 shows the followed modelling methodology based on the leave-one-outcross-validation (LOOCV) technique.

After the data division step (subchapter 4.2.1), the model design procedure, based on the LOOCV will be carried out (subchapter 4.2.2). Within this modelling procedure the model structure and hyperparameters will be updated until the best LOOCV performance is reached.

During the design and training procedure the LOOCV average error, calculated over the validation experiments (see subchapter 4.2.1) at each iteration will be used to compare the model structures. This average error is calculated using different performance metrics. The main error performance metric used in this procedure for the comparison of the model structures has been the mean absolute percentage error (MAPE) and the root mean squared error (RMSE) (Equations (15) and (17)). Nevertheless, along the process of modelling and also for the model comparison other goodness-of-fit indicators as the mean absolute error (MAE), the maximum error (MAX)

and a hit rate (HITS) are used (Equations (14) - (18)). Particularly, the HITS considers a hit whenever a model prediction leads a maximum estimation error lower than the experimental error ($0.02 \text{ g}_\circ \text{g}^{-1}$). This experimental error, which is needed in the hypothesis testing for model discrimination, was estimated by repeating several experiments at the same experimental conditions according to (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012).

$$MAX = \max_{i \in [1, \dots, n]} |f(X_i) - NN(X_i)| \quad (14)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|f(X_i) - NN(X_i)|}{f(X_i)}, \text{ where } f(X_i) \neq 0 \quad (15)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |f(X_i) - NN(X_i)| \quad (16)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(X_i) - NN(X_i))^2} \quad (17)$$

$$HITS = \#\{i \in 1, \dots, n : NN(X_i) \in [f(X_i) - \varepsilon, f(X_i) + \varepsilon]\}, \quad (18)$$

where ε is the experimental error.

These types of techniques for model building try to test the generalization capability of a model structure (Kashani & Shahhosseini 2010).

Finally, once the model structure and/or its hyperparameters have been selected and properly validated, the final model is trained with all the available data except for the test dataset which is used to estimate the final goodness-of-fit of the trained model (see subchapter 4.2.3).

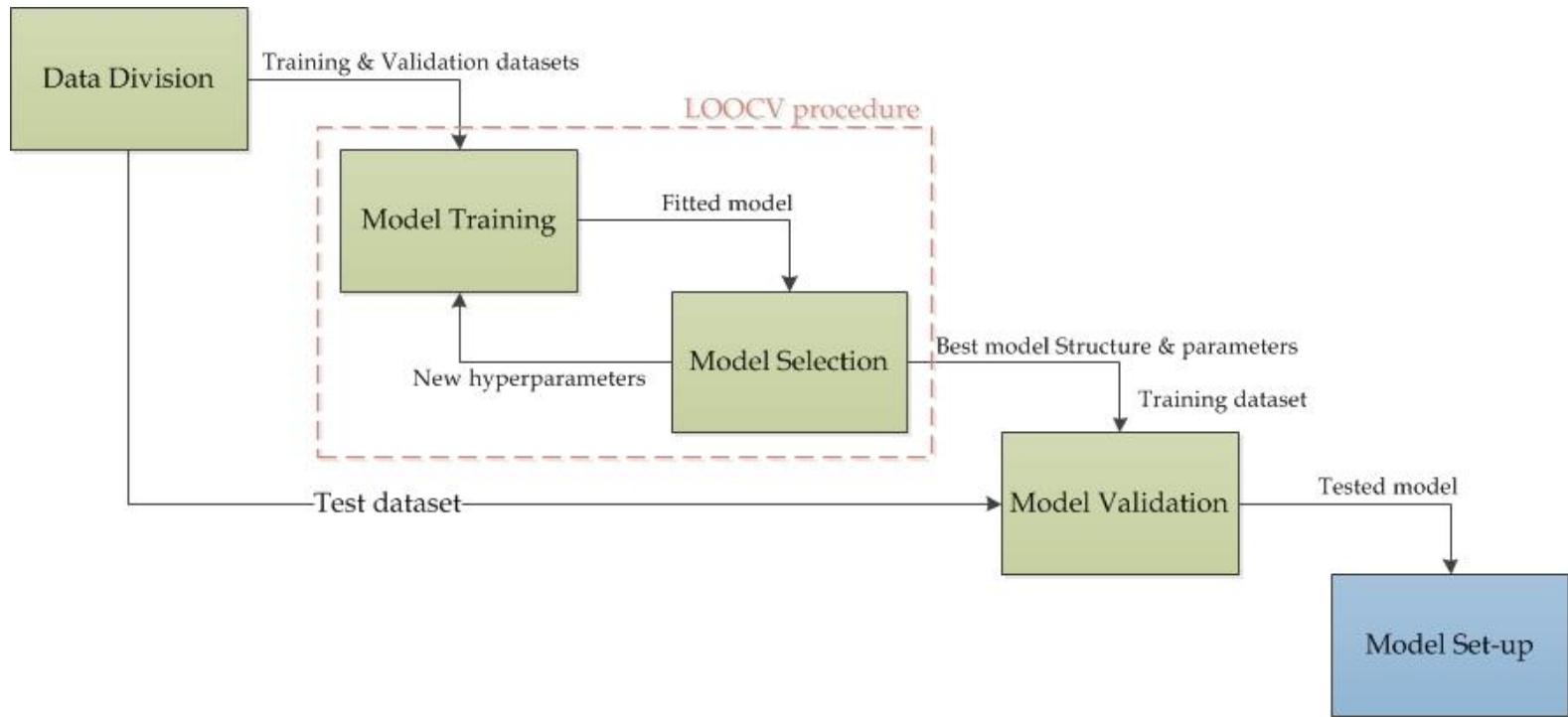


Figure 15. Scheme of the followed modelling methodology.

4.2.1. Data Division

The first step of the methodology is the data division to generate three different datasets that will be used during the modelling and validation procedure. In this point the test dataset is already separated from the training and validation datasets to keep totally unseen by the fitted models until the last model validation step. Once the data is correctly divided, an iterative procedure to select the best model structure and its hyperparameters is carried out.

Consequently, the available data has been divided into training, validation and test datasets. The training and validation datasets are used for the selection of the model structure, training algorithm, kernel functions or hyperparameters. The first one is used to train several models with different structures, using different training algorithms or kernel functions and different hyperparameters. The validation dataset consists of the excluded experiments from the training at each LOOCV iteration and it is used to calculate the validation error that is used in the early stopping concept to avoid the overfitting. The training is stopped if the validation error increases in a sufficiently high number of consecutive iterations. Finally, the test dataset will be used to obtain the goodness-of-fit of the model.

4.2.2. Model Design and Selection

As previously introduced, the model design and selection will be done following the LOOCV iterative procedure. Figure 16 shows this procedure, which is divided in the training step of each model and the resultant model selection step. For each modelling technique (see subchapter 4.1.4 Modelling techniques) several combinations of their corresponding hyperparameters with different structural models (see subchapter 4.2.2.1 Structural models) will be set up for the training procedure. Once trained, the obtained combinations of model structures and model hyperparameters will be

compared using the LOOCV average errors. Those with the best performance measurements will be selected combination.

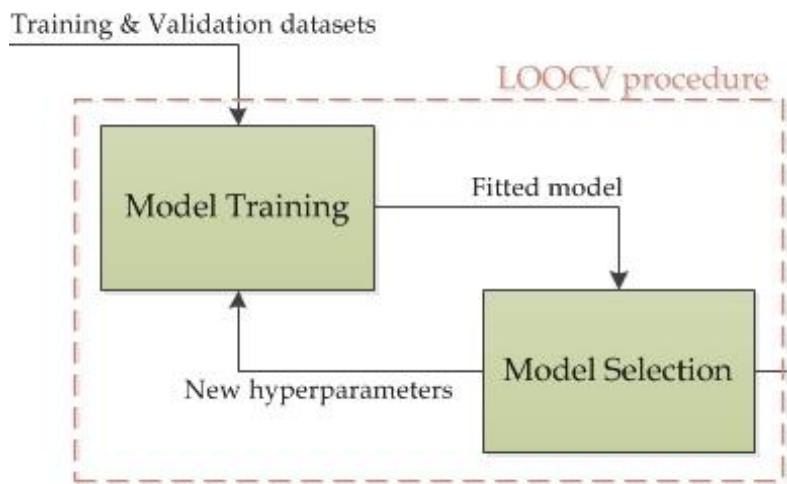


Figure 16. Scheme of the LOOCV procedure for the selection of model structure and its hyperparameters.

This iterative procedure has been carried out in order to select the model structure and parameters that better fits the process. The LOOCV technique consists of setting aside a set of experiments (representing unique operational conditions) from the model training phase and only using them for the validation phase. All the process is repeated until every single set of experiments is used in the validation stage. Figure 17 graphically represents the LOOCV training procedure. As can be seen, in each training step a different validation dataset is keep out from the training and used to calculate a LOOCV validation error. Once all individual datasets have been used for the validation the global LOOCV mean error is calculated, using this measure to compare the generalization ability of the tested structures one with each other. Therefore, the whole iterative training procedure is repeated modifying the structure or the hyperparameters of the proposed model until the best model structure and parameters are found. Note that the test dataset is keep unseen by the modelling technique in the whole LOOCV procedure.

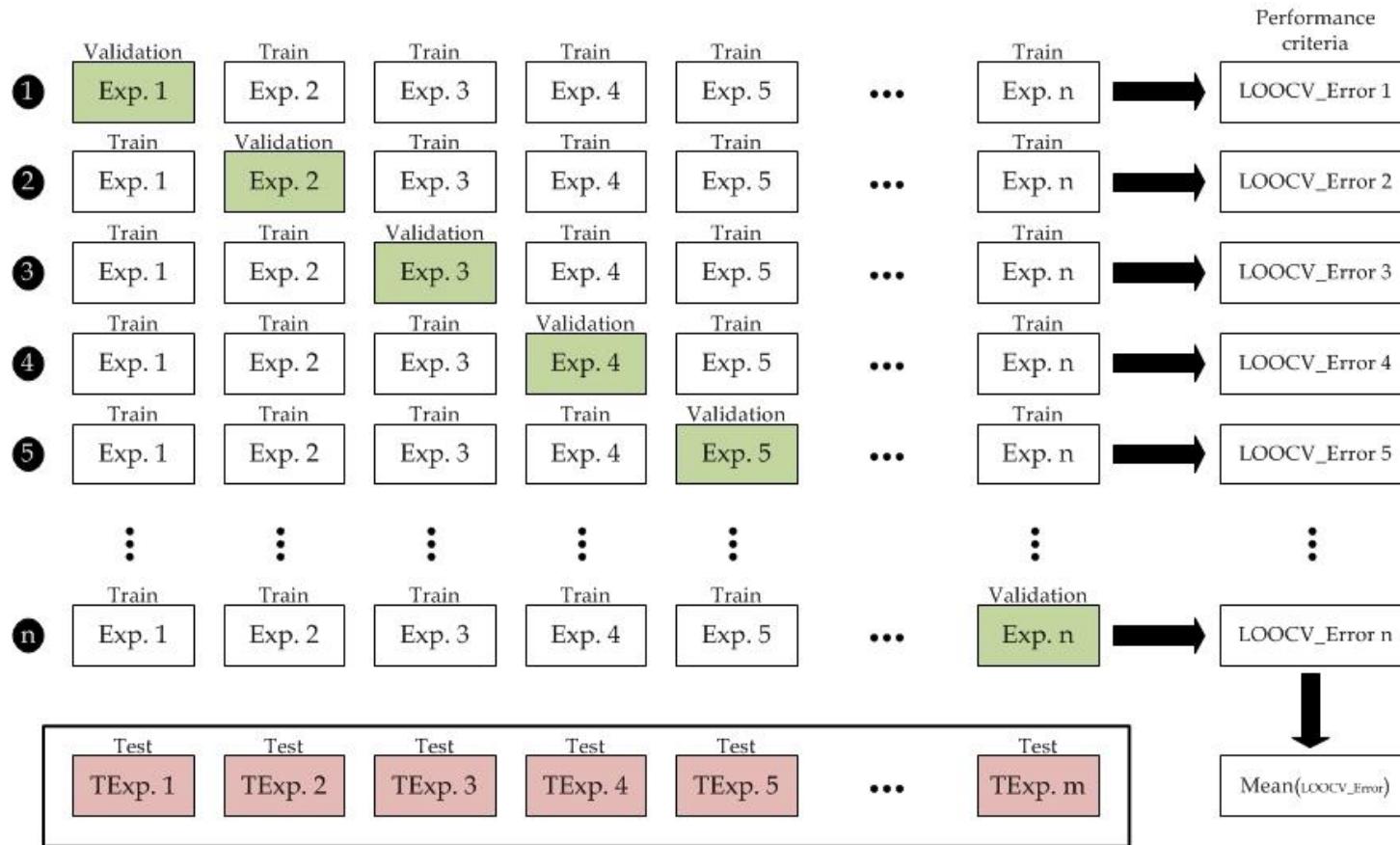


Figure 17. Diagram of the LOOCV training procedure

4.2.2.1. Structural models

In addition to the internal topology, structure and/or hyperparameters of the models, which will depend on the selected modelling technique, another point that should be taken into account in the modelling procedure is the external structural model. It has been to be selected which information will dispose the model to estimate the desired outputs. It must be defined the input-outputs vectors of the models.

Therefore, all the LOOCV based modelling methodology has been carried out with different input-output structures for each modelling technique and training parameters. Next, the main structural models extracted from the literature are presented. However, several variants of some of them have been also tested during the modelling procedure.

- FIR (Finite Impulse Response): It is the simplest model. The estimations depend only on previous input variables.

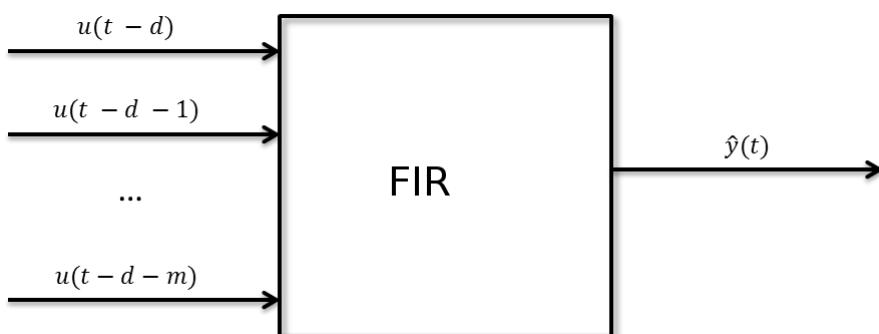


Figure 18. FIR structural model scheme

- ARX (AutoRegressive with eXogeneous inputs): The estimations of these type of models depends from previous both inputs and outputs of the process. A variant of this structure are the Non-Linear AutoRegressive models with eXogeneous inputs (NARX). These structures have obtained very good results identifying non-linear systems.

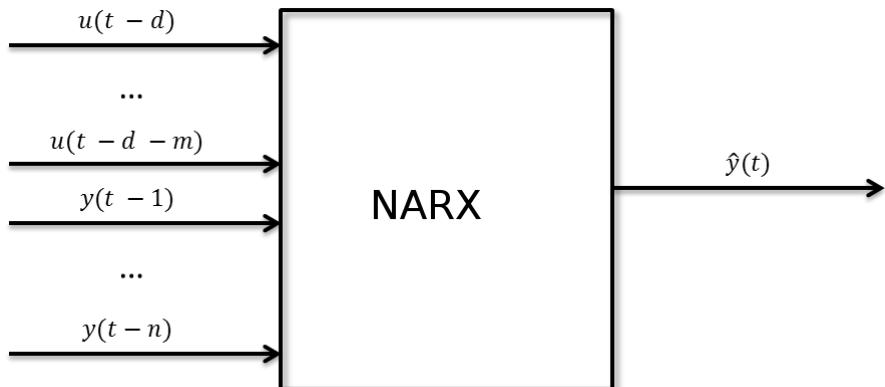


Figure 19. NARX structural model scheme

These two structural models are the most used ones with satisfactory results (Meert & Rijckaert 1998; Molga 2003; Mujtaba et al. 2006; Rahimpour et al. 2011). Therefore, several configurations of these structures have been implemented and tested. Table 3 shows the best input-output vector configurations found during the modelling procedure. The first four ones and modifications of them were initially implemented with satisfactory results. However, later it was observed that for long-term production estimations it is necessary to provide the model with information about the catalyst deactivation. Note that the information about the activity data is not available because is a non-measurable data and therefore initially was not included as a possible input. Thus the activity was included as new inputs of a neural network implemented in an hybrid neural model (see subchapter 5.1.4).

Moreover to the classification between the FIR and NARX structures, some of these configurations have more or less input/output memory in order to improve the capability of the model to assimilate the dynamic behaviour of the process. Other configurations includes more information about the process such the time vector or the activity level of the catalyst and its past evolution.

Table 3. Examples of some of the most important tested input-output configurations.

| Type | Input Vector | Output Vector |
|------|--|---------------|
| FIR | $X_w(k), T(k), W/F_{E0}(k)$ | $X_O(k)$ |
| FIR | $X_w(k-1), X_w(k), T(k-1), T(k), W/F_{E0}(k-1), W/F_{E0}(k)$ | $X_O(k)$ |
| NARX | $X_w(k-1), X_w(k), T(k-1), T(k), W/F_{E0}(k-1), W/F_{E0}(k), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k-1), X_w(k), T(k-1), T(k), W/F_{E0}(k-1), W/F_{E0}(k), t(k-2), t(k-1), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k), T(k), W/F_{E0}(k), X_O(k-3), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k-1), X_w(k), T(k-2), T(k-1), T(k), W/F_{E0}(k-1), W/F_{E0}(k), \Delta t(k-1), \Delta t(k), a(k-1), a(k), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k-1), X_w(k), T(k-1), T(k), W/F_{E0}(k-1), W/F_{E0}(k), a(k-1), a(k), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k), T(k), W/F_{E0}(k), \bar{X}_O(k-3), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k), T(k-1), T(k), W/F_{E0}(k), t(k-1), t(k), a(k), X_O(k-3), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| NARX | $X_w(k), T(k), W/F_{E0}(k), a(k-3), a(k-2), a(k-1), a(k), X_O(k-4), X_O(k-3), X_O(k-2), X_O(k-1)$ | $X_O(k)$ |
| FIR | $X_w(k), T(k), W/F_{E0}(k), a(k-3), a(k-2), a(k-1), a(k)$ | $X_O(k)$ |

Other important structural models used in the literature are the ARMAX, SSIF type models. In both cases it is necessary to know the target variable in order to introduce into the model the current deviation. Therefore, for long-term estimation procedures or for offline optimisation purposes these models cannot be used. However, they have good features for control purposes and online optimisation procedures.

- ARMAX (AutoRegressive Moving Average with eXogeneous inputs): One of the most complete structural models, as ARX models, their estimates from previous inputs and outputs of the system and they incorporate as model inputs the estimation error committed in some previous steps.

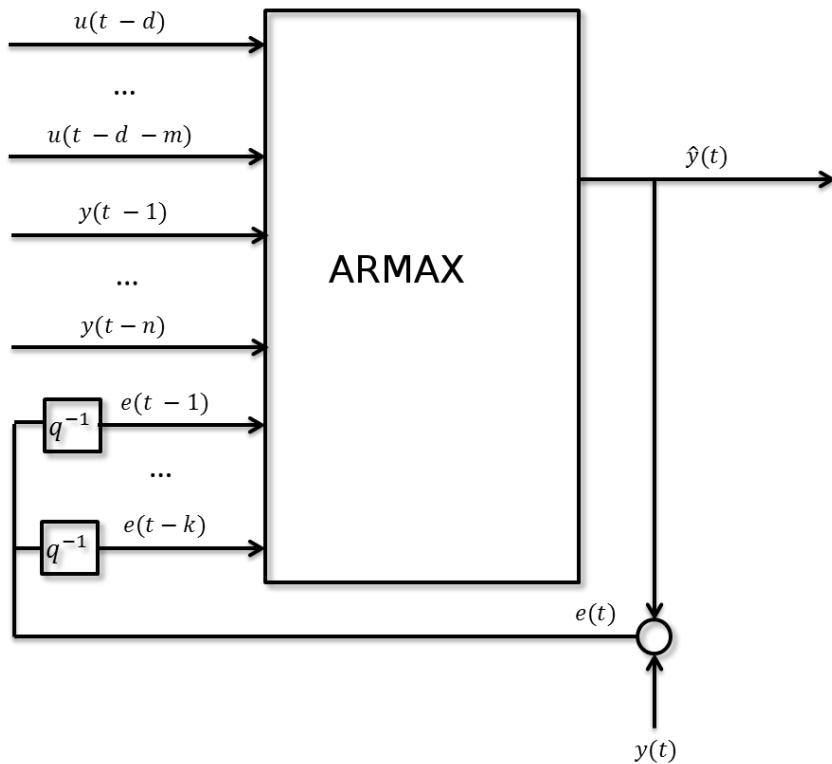


Figure 20. ARMAX structural model scheme

- SSIF (State Space Innovations Form): This model type is completely different from the previous ones. This one works in the state space and it is an alternative to the classical input/output characterization.

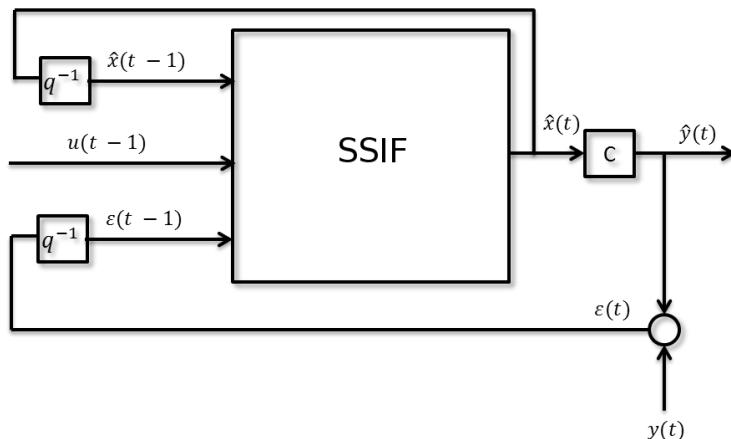


Figure 21. SSIF structural model scheme

4.2.2.2 Model Design and Comparison

As previously presented, during the modelling stage, and using the LOOCV procedure, several combinations of modelling technique, model hyperparameters and structural models will be tested. Figure 22 shows the modelling techniques that will be tested for their capability for both description and prediction of the behaviour of the BTO process. The selected model must be able to predict with the desired accuracy in both situations. Otherwise it will be rejected. If several combinations meet the requirements, following the parsimonious method, the solution with the least complexity and requiring the least number of variable and hyperparameters will be selected.

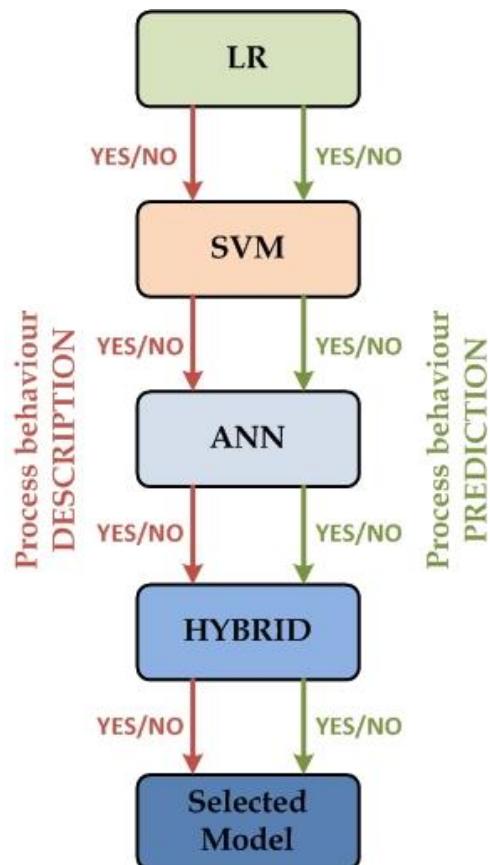


Figure 22. Model design and selection procedure based on prediction capability

Linear Regression (LR) based modelling technique, will be consider as the most simple and with lower computational cost solution for the modelling problem. Next the SVM and ANN will be trained and tested following the soft-modelling methodology. And finally, if it is necessary, a hybrid modelling approach will be implemented with the most promising modelling technique (SVM or ANN).

Finally, the obtained models with each modelling technique will be compared and analysed by their capability to predict the BTO process. For this purposes, a synthetic test dataset generated by the mechanistic model will be used. This new and unseen test dataset includes both complex dynamic trajectories and long-term experimental conditions such as the ones that will appear in reality.

4.2.3. Model Validation

Once the most suitable model structure and hyperparameters have been selected following the cross-validation procedure, the final model should be validated. Figure 23 shows the final step of the followed modelling methodology. The final model with the selected model structure and hyperparameters is trained with the completed training dataset and validated with the already unseen test dataset in order to test its generalization capability.

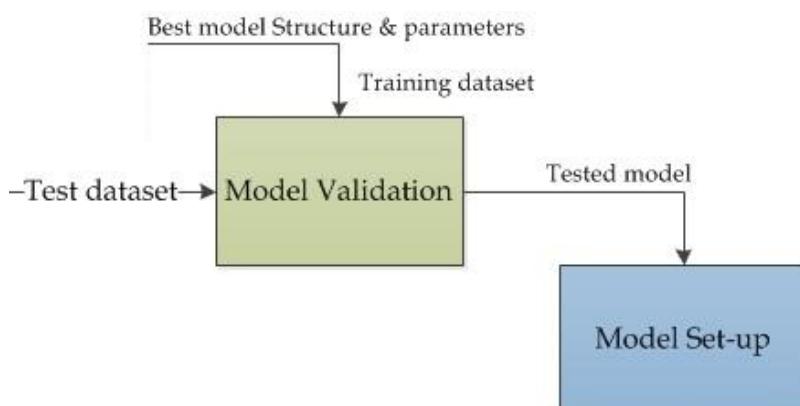


Figure 23. Final validation of the selected models

As previously exposed, the performance metrics exposed in the Equations (14) - (18) will be used. Finally, the mechanistic model implemented following the kinetic scheme proposed by Gayubo et al. (2010; 2012) will be used as contrast method. Table 4 presents the error analysis for this model and therefore the values that will be used as modelling objective. The obtained results in the modelling procedure by the new models will be compared with the results provided by this mathematical model. Additionally, the coefficient of determination (R^2) between the original data and the estimated data will be also used with an equivalence analysis that will be performed in order to completely validate de proposed soft-modelling approach. And finally, an equivalence test will be used to compare the final fitted model with both the well-known mechanistic model and with the real experimental data. To this end, the Two One-Sided Test (TOST) approach has been selected (Robinson et al. 2005).

Table 4. Error analysis of MECH model

| | MAE (g _o g-1) | RMSE (g _o g-1) | MAPE (%) |
|------|--------------------------|---------------------------|----------|
| MECH | 0.0263 | 0.0323 | 15.7886 |

4.3. Optimisation methodology

The second objective is to identify the best operation policy to maximise the total production. The current operation of chemical processes use to be limited to constant control set-points optimised at zero time of stream. This simple strategy is the optimal one for simpler processes and for those processes where their catalyst is maintain stable and do not significantly deactivate. However, for processes as the BTO where the catalyst deactivation significantly influences into the final production or if the catalyst regeneration costs are very high, the optimal operation strategy should take into account the catalyst deactivation.

In the present work the process behaviour, taking into account the effect of catalyst deactivation, is simulated under each proposed operational conditions within the optimisation procedure. The aim of the optimisation problem is to obtain optimal operational conditions for the BTO process in order to maximize the Total Production of Olefins per Space-Time (TPOST). In this way, the olefins production, given by the Equation (19), defines the objective function to be maximized.

$$\max(TPOST) = \max_{T, X_W, WF_{E0}^{-1}} \frac{\int_0^{t_f} R_o(T, X_W, WF_{E0}^{-1}) dt}{WF_{E0}^{-1}} \quad (19)$$

The term R_o represents the yield of olefins per equivalent ethylene feed. Please note that the stopping criteria and therefore the definition of t_f constant in the upper bound of the integral can be reached by two ways:

- The activity is not below 0.10. The catalyst involves a major process cost; therefore to maximize the production per amount of catalyst we have set a lower bound on the activity in order to reduce total production costs.
- The olefins conversion rate does not decrease below 0.10. This ensures that a stable production is maintained.

If any of the above criteria is violated, it is considered that the production has reached its maximum span and should be stopped to proceed with a catalyst regeneration phase. Moreover, the following variables are the main drivers of the process behaviour. These operating variables are the optimising variables:

- Operating variables:

T: reaction temperature (K).

X_W : mass fraction of water based on the equivalent mass of ethylene in the reactor feed, $\text{g}_{\text{water}}\text{g}^{-1}$

$$WF^{-1}_{EO}: \text{space-time, } g_{\text{catalyst}} h^{-1}(g_{\text{ethanol}})$$

- Activity level:

a: catalyst activity. The activity has been considered as a disturbance that quantifies the rate of catalyst deactivation by coke.

The temperature is the dominant operating variable of the process. For this reason three different strategies have been proposed and compared to generate the temperature profiles in the first and second strategies, and the dynamic profiles for each operating variable in the third option. These three proposed optimisation strategies use the Equation (19) as fitness function. Namely, the process is simulated under the conditions given by each solution until one of the termination condition happens and retrieving the total olefins production (TPOST).

Therefore, during the optimisation procedure it is necessary to simulate the process behaviour hundreds of times with each different combination of the optimisation parameters. Due to this necessity of performing a huge quantity of simulations of the BTO process, these have to be done with a model with a limited computational cost. Otherwise each optimisation procedure could be prolonged for months, delaying possible future developments.

The use of surrogated models for the simulation of the process is therefore necessary. In the case of the BTO process, its simulation during the optimisation procedure will be carried out using the soft computing based models previously explained. Due to their characteristics, these modelling approaches present significantly less computational costs. However, in order to fully assess that the obtained results optimising the operational conditions with a black-box type model to guide the procedure are correct without running the corresponding experimental runs, the whole optimisation procedure will be carried out using the previously explained mechanistic

(MECH) model. This model is a well-known model of the BTO process and is recognized by the scientific community.

4.3.1. Constant optimisation

The first optimisation strategy is to search for the best constant set-point combination that maximizes the total production of olefins (TPOST). This strategy is commonly used in the industry and in this work has been used, with the optimisation at zero time of stream (see subchapter 4.3.4), as control method for comparing purposes. It is assumed the hypothesis that a dynamic optimisation of the operational conditions will obtain significantly better results than maintaining constant set-points, although they were optimised. Therefore the constant set-points (Equation (20)) will be used for contrast purposes.

$$f_0(t) = \beta \quad (20)$$

4.3.2. Dynamic optimisation

The second strategy is focused on considering different trajectory types for the temperature and their effect over the production of olefins. Several trajectories have been proposed for the temperature. The most representative ones can be seen in Equations (21) - (29).

$$f_1(t) = \alpha t + \beta, \quad (21)$$

$$f_2(t) = \alpha e^{\gamma t} + \beta, \quad (22)$$

$$f_3(t) = \alpha \tan^{-1}(\gamma t) + \beta, \quad (23)$$

$$f_4(t) = \alpha \ln(\gamma t + 1) + \beta, \quad (24)$$

$$f_5(t) = \alpha t^2 + \beta, \quad (25)$$

$$f_6(t) = \alpha t^3 + \beta, \quad (26)$$

$$f_7(t) = \alpha a^{-1} + \beta, \quad (27)$$

$$f_8(t) = \alpha a^{-2} + \beta, \quad (28)$$

$$f_9(t) = \alpha a^{-3} + \beta. \quad (29)$$

The above trajectories have up to three parameters that should be optimised: α , β and γ . These trajectories drive the process temperature set-point until the maximum admissible temperature of the process is reached. The parameter β always represents the intercept of the above trajectories, being as consequence their feasible values depending on the physical-chemical properties of the process. In fact, in this case, 573 K is the inferior bound where the complete dehydration of the ethanol happens and 673 K is the upper bound to avoid the irreversible deactivation of the catalyst.

On the other hand, α and γ could be freely adjusted. Preliminary results suggest that α and γ always reach their optimum values in the interval [0; 5] so in this work the [0; 10] interval has been set.

Furthermore, the variables X_W and WF^{-1}_{E0} will always being considered as constants and will be also optimised. In particular, X_W will range between [0.0821; 4.8889] and WF^{-1}_{E0} will range between [0.068; 1.525] respectively. Please note that those intervals have been chosen as the ones used to adjust the mechanistic model (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012).

The generated solutions of each iteration within the optimisation procedure are real vectors containing X_W and WF^{-1}_{E0} operating variables and the parameters of the proposed temperature trajectories.

4.3.3. Dynamic neural optimisation

The third optimisation strategy consists of the use of Artificial Neural Networks to directly generate a dynamic operating profile for each of the operating variables and not only for the temperature. This third strategy has been chosen due to the complexity (14^n) of testing each possible combination of the proposed trajectory profiles (see subchapter 4.3.2) for the three main operating variables.

The optimisation procedure will be used in this case to directly train the weights matrices and bias vectors of an ANN with the topology shown in Figure 24. Therefore, in this case, the generated solutions of the optimisation procedure are formed by real vectors containing the weights and biases of the ANN that generates the dynamic trajectories of each operating variable. During the optimisation procedure, these matrices and vectors defining the mapping computed by the ANN are updated in order to define the ANN that generates each control trajectory that maximizes the TPOST. This ANN receives as inputs the operational conditions of the process (T , X_w , $WF^{1_{E0}}$) for the previous time step and the current catalyst activity level (a). The output of the network is the current set point for each variable. Different neural structures have been trained and tested with one hidden layer and with different number of neurons in it (from 4 to 8). In this point the recommendations previously stated in the literature have been followed (Molga 2003), testing up to $2N$ neurons in the hidden layer, being N the number of inputs of the neural network.

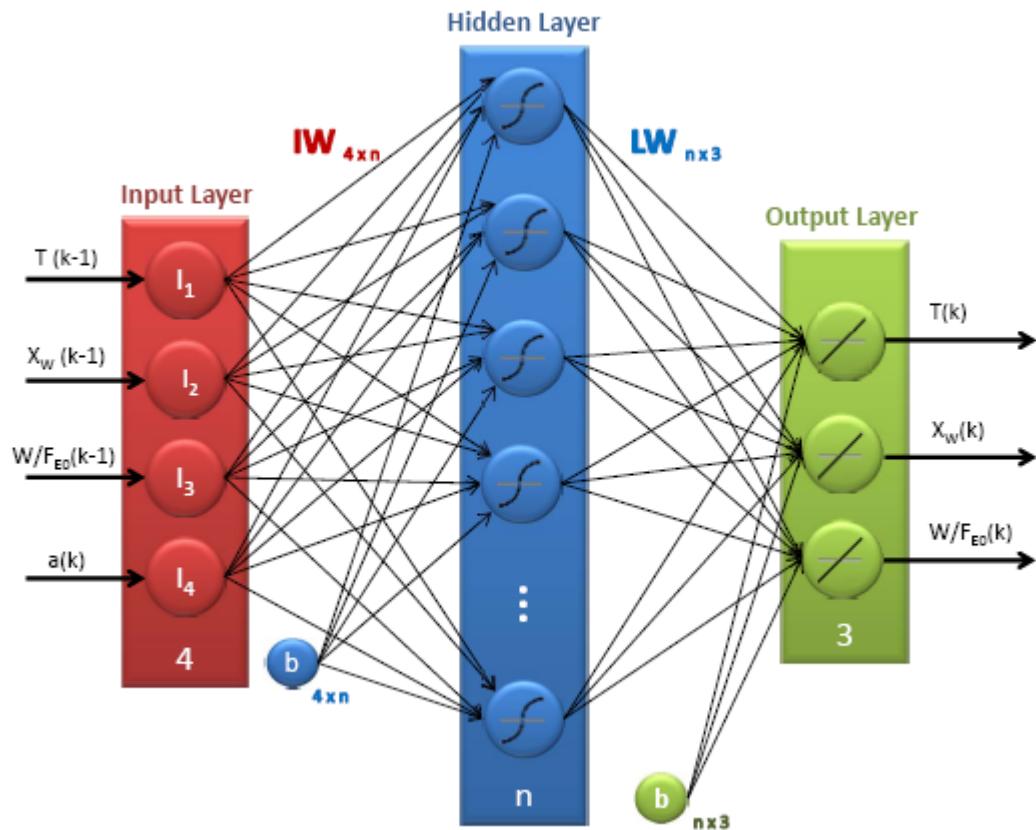


Figure 24. Feed-Forward ANN topology for generating the trajectories of each operating variable.

4.3.4. Optimisation at Zero Time of Stream

The study of the processes at zero time of stream is the standard optimisation procedure to study the operational conditions for catalytic processes. Thus, this procedure will be used as a contrast method for the proposed optimisation strategies.

In contrast to the proposed optimisation strategies, this option does not take into account the catalyst deactivation rate. The optimisation consist on searching for the best constant operational conditions that maximise the conversion rate for olefins at zero time on stream, independently to the behaviour of the process under the selected set-points.

The cost function, given by the Equation (30), defines the new objective function to be optimised in this optimisation procedure; which consist on the TPOST at zero time of stream ($a = 1$).

$$\max(TPOST_{a=1}) = \max_{T, X_W, WF_{E0}^{-1}} \left(\frac{R_O(T, X_W, WF_{E0}^{-1})}{WF_{E0}^{-1}} \right)_{a=1} \quad (30)$$

The present work is on the hypothesis that disregarding the catalyst deactivation causes the resulting operational conditions will deactivate it too early decreasing the total production and increasing the catalyst regeneration costs.

4.3.5. Evolutionary optimisation

These three presented optimisation strategies, and the current optimisation method at zero time of stream, have been implemented and solve using the same optimisation technique. Due to the incremental complexity of the problem, from a simple optimisation problem (constant set-points) to a numerically complex problem with almost a hundred of optimising variables (dynamic neural optimisation), the evolutionary algorithms have been selected to resolve the optimisation problem.

Each individual of the evolutionary algorithm is a solution for the corresponding optimisation strategy. And each evaluation of an individual implies the simulation of the complete behaviour of the process under the proposed solution.

One of the problems of the evolutionary optimisation techniques is the computational cost. During the searching procedure the fitness functions of the problem will be run hundreds of times. Thus, when the computational cost of the fitness function is quite high, the evolutionary optimisation can be unable to resolve the problem in acceptable times.

Therefore, it is commonly used in the literature the use of surrogate models of the processes being optimised. Surrogate models based on soft-computing techniques are able to maintain an acceptable equivalence with the real objective but with much less computational cost. Thus, they can be integrated in the fitness function of the optimisation problem to simulate the behaviour of the process with a reduced error. In this work the previously presented neural model will be used as surrogate model of the BTO process.

4.3.5.1. Evolutionary Algorithms

The Evolutionary Algorithms (EA) are soft computing techniques that with their robust and adaptability capabilities are used to solve non-linear, highly dimensional and complex engineering problems (Oduguwa et al. 2005). They form a branch of artificial intelligence especially dedicated to solve optimisation problems. Inspired in the biological evolution process, the EA are able to adapt to the changing nature of the reality finding solutions that standard optimisation techniques are unable to reach.

The main EA can be classified in the families show in Figure 25. Being the Genetic Algorithms (GA) the first one to be developed, each group has been oriented to solve more specific optimisation problems.

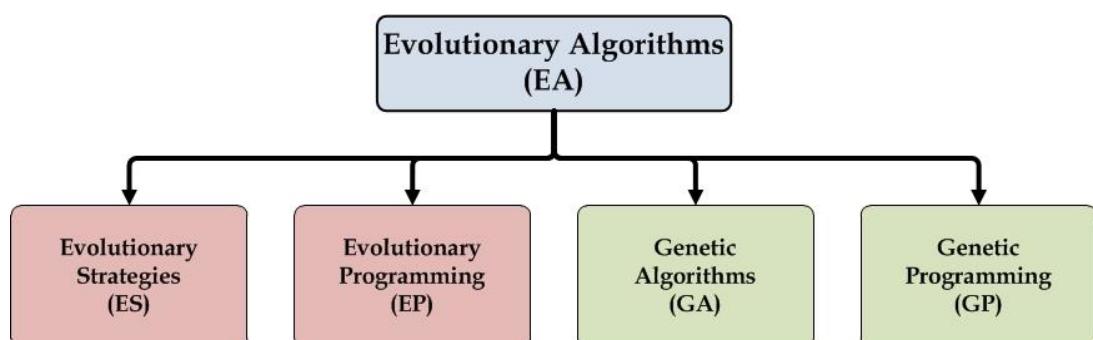


Figure 25. Classification families of the Evolutionary Algorithms.

All of them have some basic characteristics in common. Figure 26 shows these common concepts present in all EA and which are inspired in the genetic evolution processes of the biological organism:

- **Population:** the population is composed by a number of possible solutions or **Individuals** for the optimisation problem. The initial population used to be randomly generated with different **genesis operators**.
- **Individuals:** each individual of the population represents a unique codified solution to the optimisation problem. Each of these candidates can be encoded as a real, integer or binary values of vectors or matrices, also called **Chromosomes**.
- **Fitness evaluation:** the individuals are evaluated using the defined fitness function of the optimisation problem to establish the **aptitude** of each one. Based on this fitness value, the **reproduction** possibilities (selection as parents) may be different.
- **Parent Selection:** from the total population using different selection operators and strategies a number of **Parents** are selected to take part in the evolution to the next generation. Depending on the evolution strategy or the previously presented EA family type, all individuals take parts in the evolution or only the selected ones.
- **Crossover / Mutation:** the previously selected parents will be modified using the recombination operators of **Crossover** and **Mutation** in order to generate new **Offsprings**. Crossover operators will exchange genetic material between two or more parents to generate one or more new offsprings. The mutation operator will produce a modification in the genetic material of a specific chromosome to generate the new offsprings. Usually, the mutation operator is needed to prevent a premature convergence of the algorithm in local optimums, randomly exploring new search areas.

- **Generational replacement:** finally, the initial population is replaced by a new one, composed of the evolved individuals (offsprings and also, in some cases, parents). Depending on the EA type and additional operators, such as the elitism, a complete new population replace the whole initial population or only part of it. For example, using the **Elitism operator**, one or more of the best solutions (parents or offsprings) of each generation is maintained from one generation to the following one. In addition, some algorithms maintain stable the size of the population while other ones modified it depending on their evolution strategy.
- **Restrictions:** in all the previous evolution steps, the restrictions coming from the optimisation problem can be included in order to limit or restrict the possible searching space or valid solutions (individuales).

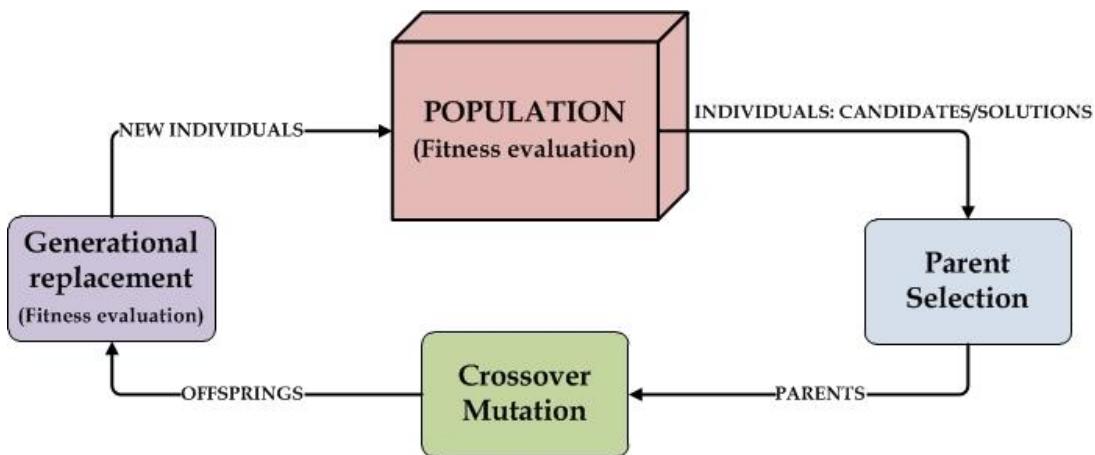


Figure 26. Basic principles of the Evolutionary Algorithms.

Although all the EA families previously presented have in common these principles, each of them applies them in a different way.

The GA were developed by John Holland in 1975 at the University of Michigan (Holland 1975) to solve search and optimisation problems. This

type of EA generates the new individuals using the crossover and mutation operators and creates the population for the next generation on the basis of the fitness of each individual.

The GP uses the same principles used by the GA but applied to the evolution of computer programs. The main difference is the codification of each individual to represent a solution. The computer programs are represented using tree structures and each solution represents specific branches and tree nodes.

The ES groups several of optimisation algorithms using the same evolution principles of crossover and mutation but following different strategies to represent the population, each individual or to perform the generational replacement. In this type of algorithms is common that each individual generates an offspring by mutation. A successful example of these types of algorithms is the Covariance Matrix Adaptive - Evolution Strategy (CMA-ES) and several of its variants, which have been successfully used in several real-world applications. The CMA-ES uses the adaptation of the covariance matrix of the selected best individuals to generate the new population.

Finally, the EP are similar algorithms to the ES but their evolution capabilities are oriented to the mutation procedures with uniform probability and often with adaptive strategies. The populations for each new generation are selected using a combination of n best parents and m best offsprings.

In the present work, the GA have been selected to solve the exposed optimisation problem. Table 5 summarizes the principal parameters of the EA implemented. First, a random initial population of 50 individuals is generated. The generated individuals are real vectors containing the optimizing variables. These are the constant operation set-points for the first optimisation strategy (see subchapter 4.3.1); X_w and WF^{-1}_{E0} operating variables and the parameters of the proposed temperature trajectories for the second optimisation strategy (see subchapter 4.3.2); or real vectors

containing the weights and biases of the ANN that generates the dynamic trajectories of each operating variable (see subchapter 4.3.3). Each individual of the population is evaluated using the MECH model or a surrogated model of the BTO process to simulate the whole process and estimate the TPOST (Equation (19)).

The 4-Tournament procedure has been chosen as selection operator of the EA. After the selection of the parents, the crossover and mutation operators are used to create the next generation. For the crossover operator the arithmetic crossover has been chosen and the adaptive operator, meeting the existing bounds and constraints, as mutation operator. Finally, following the Elitism criteria the best solution of each generation is maintain in the new generation.

Table 5. Main parameters of the EA used to optimise the trajectories.

| Parameters | Value |
|------------------------------|------------------------|
| Individuals | Vector of real numbers |
| Population | 50 individuals |
| Generations | 100 generations |
| Fitness Operator | Olefins Production |
| Genesis Operator | Random initialization |
| Selection Operator | 4-Tournament |
| Crossover Operator | Arithmetic crossover |
| Mutation Operator | Adaptive mutation |
| Crossover Probability | 0.8 |
| Elitism | True |

MODELLING OF THE BTO PROCESS

In this chapter the modelling of the BTO process using several soft computing techniques is detailed. The novelty of the BTO process, whose complete behaviour is still currently under investigation, difficult the modelling using these type of techniques. However, the optimisation and future scalability to industrial scale reactors of this process requires from surrogated models that accelerate the development of advanced and optimized control strategies for this and other biorefinery processes.

In the specific case of the BTO process, the resulting model should adequately reproduce the dynamics of the process, which is the result of both the kinetics of the process itself, and the deactivation kinetics of the catalyst which promotes the reaction. Therefore, it must be taken into account the two behaviours simultaneously.

The model design and training procedure for global models, using the LR, SVM and ANN modelling techniques, has been performed using the experimental data (see section 4.1.2 Experimental Data). Instead, the hybrid approach has been carried out using the synthetic data (see section 4.1.3 Synthetic Data) to provide to the model with dynamic information about the catalyst deactivation.

5.1. Modelling results

5.1.1. Linear Regression

The Linear Regression (LR) modelling procedure has been selected for contrast purposes. The objective has been to test if a linear regression based modelling technique was enough to assimilate the behaviour of the BTO process.

The Regression toolbox of the MATLAB® software has been used to adjust different linear models with the previously presented model structures: BASE, DYN, REC and TEMP. Two different algorithms have been used to adjust the regression parameters of the linear models: the QR decomposition method and the least absolute shrinkage and selection operator (LASSO) regression method.

Table 6 shows the average estimation error obtained during the LOOCV procedure for all the real experiments. The linear model fitted with the QR decomposition method shows significantly better results than the LASSO regression algorithm. Some models as the BASE and DYN models presents similar performance compared with the SVM based models (see subchapter 5.1.2). However, they show poorer results than the results obtained with the ANN (see subchapter 5.1.3).

Table 6. Estimation average errors with the LOOCV procedure for all real experiments with LR based models

| | QR decomposition | | LASSO regression | |
|-------------|-----------------------|-------------|-----------------------|-------------|
| | RMSE (golefinsg-1) | MAPE (%) | RMSE (golefinsg-1) | MAPE (%) |
| BASE | 0,0374 | 26,2407 | 0,1860 | 120,713 |
| DYN | 0,0373 | 25,2303 | 0,1870 | 118,578 |
| REC | 0,0494 | 39,0553 | 0,1346 | 72,2540 |
| TEMP | 0,0517 | 41,5084 | 0,1210 | 64,8448 |

Figure 27 and Figure 28 shows the process performance under two real experimental conditions with a constant temperature experiment and a variable temperature experiment respectively. In these examples and for all experimental data the BASE and DYN models fit better to the experimental data than the REC and TEMP structures, especially for variable temperature experiments.

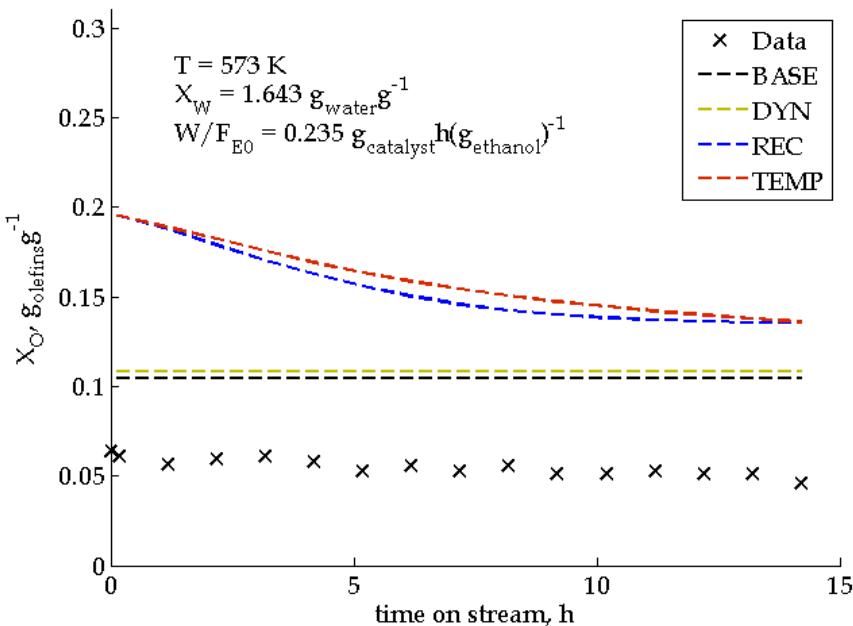


Figure 27. Example of the LR based model performance for a constant temperature experiment

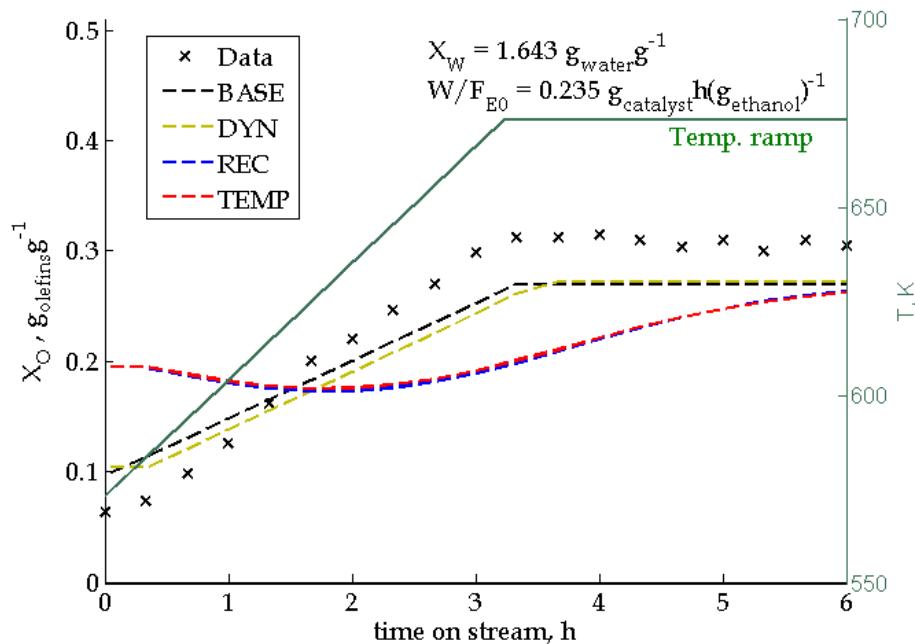


Figure 28. Example of the LR based model performance for a variable temperature experiment

5.1.2. Support Vector Regression

In order to define the most suitable SVR based model, the previously presented methodology has been applied. A grid search of the main SVR parameters has been performed in order to select the hyperparameters that provide the best estimates. The hyperparameter selection has been carried out following the LOOCV modelling technique previously explained. All the available data for training and validation (but not the test dataset which is maintained unseen by the model during the training procedure) have been used during the structure selection procedure to select the best combination of hyperparameters.

Figure 29-Figure 31 show the variation of the maximum estimation error (MAX) generated based on the main SVR parameters. These parameters are gamma (γ), cost (C) and nu (v). The parameters are faced to each other in order to draw in X-Y graphs to identify the parameter areas that minimizes the MAX estimation error. These areas are refined with a new search using a finer mesh in order to identify the best solutions. It is observed that the best hyperparameter values are located in ranges with low values near the left-bottom corner of the figures, between 1 and 7.5 for C, 1 and 3.25 for v and 1 and 6 for γ approximately.

By checking the numerical results the best solutions of γ , C and v are around 0.01, 1 and 0.02 respectively. These values have been used for the SVM based modelling of the BTO process.

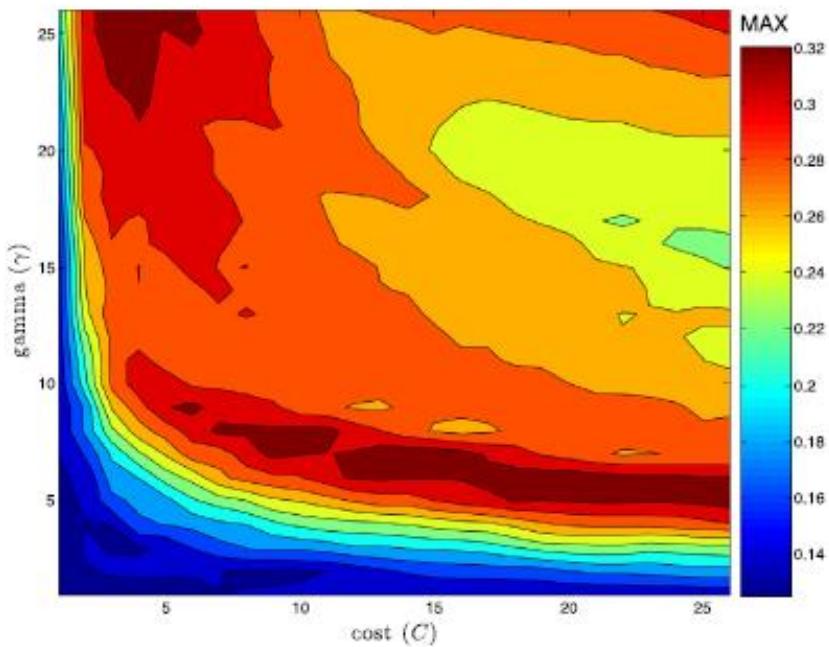


Figure 29. Effect of the c and γ parameters on the SVM based model estimates

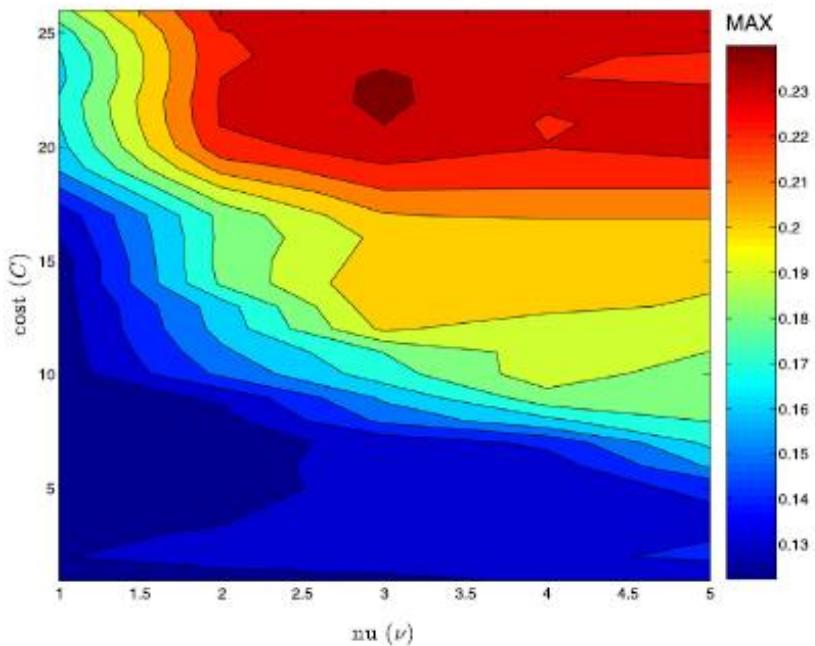


Figure 30. Effect of the c and ν parameters on the SVM based model estimates

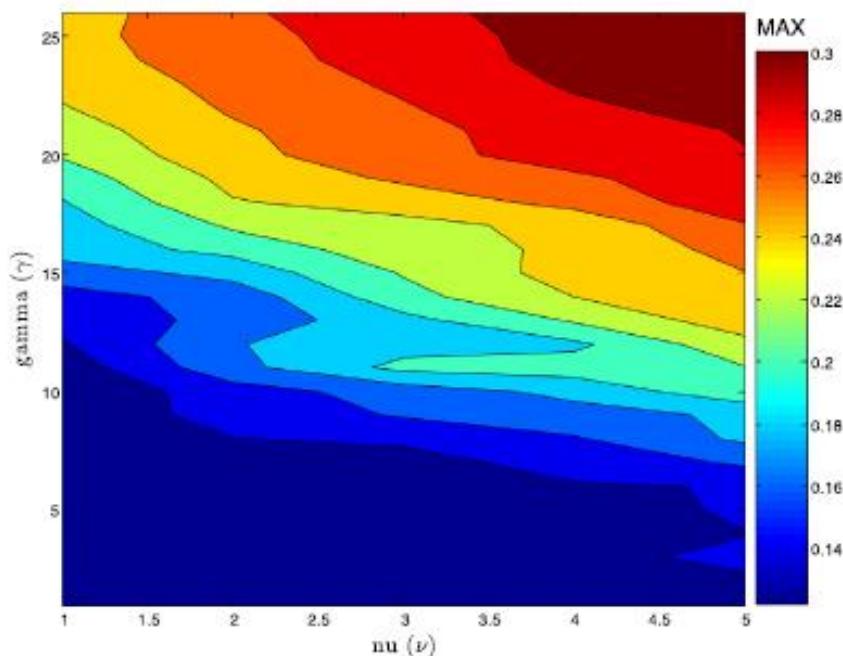


Figure 31. Effect of the γ and ν parameters on the SVM based model estimates

In Table 7, the mean estimation results (average of all the experiments) obtained for SVR based models are shown. The SVR model presents a mean error of 15.06 percentage points higher than the MECH model in terms of MAPE when using both the constant and ramp temperature profiles.

Table 7. Estimation mean errors obtained for all real experiments with SVR based model

| | RMSE (golefins g^{-1}) | MAE (golefins g^{-1}) | MAPE (%) | HITS (%) |
|--------------------|------------------------------|-----------------------------|-------------|-------------|
| T _{const} | 0.0349 | 0.0321 | 21.2654 | 28.9726 |
| T _{ramp} | 0.0432 | 0.0376 | 28.2638 | 28.4158 |
| Total | 0.0382 | 0.0342 | 24.0039 | 28.7182 |

Note that these important differences in terms of MAPE are not so representative in terms of RMSE or MAE since the absolute values of olefin ratio conversions have low absolute values. However, the differences in terms of the goodness-of-fit between SVR based models and the MECH

model is high enough to discard this type of technique to model the BTO process. Figure 32 and Figure 33 show two examples of the estimation capability of SVR based model for both constant and ramp experiments. As can be seen, this model is unable to correctly follow neither the experimental data nor the data generated by the MECH model.

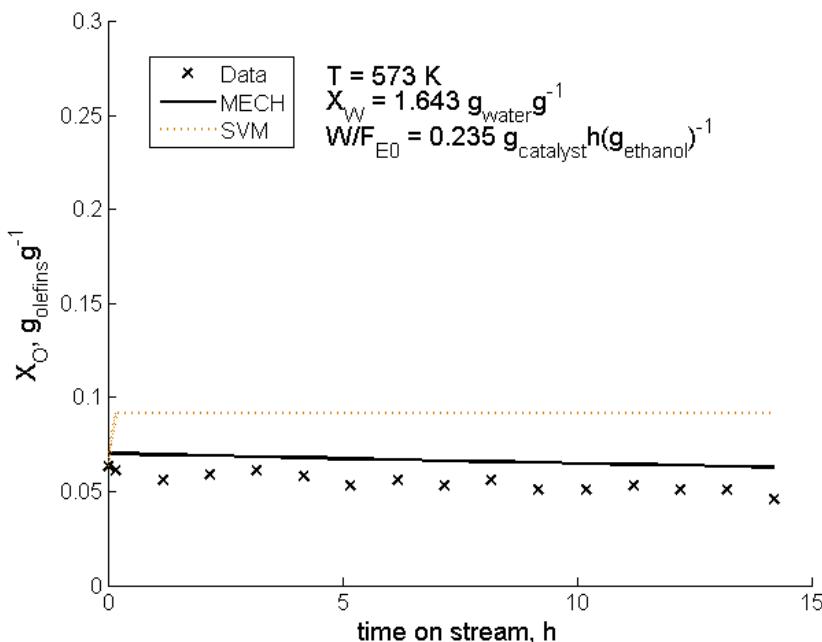


Figure 32. Example of the SVR based model performance for a constant temperature experiment

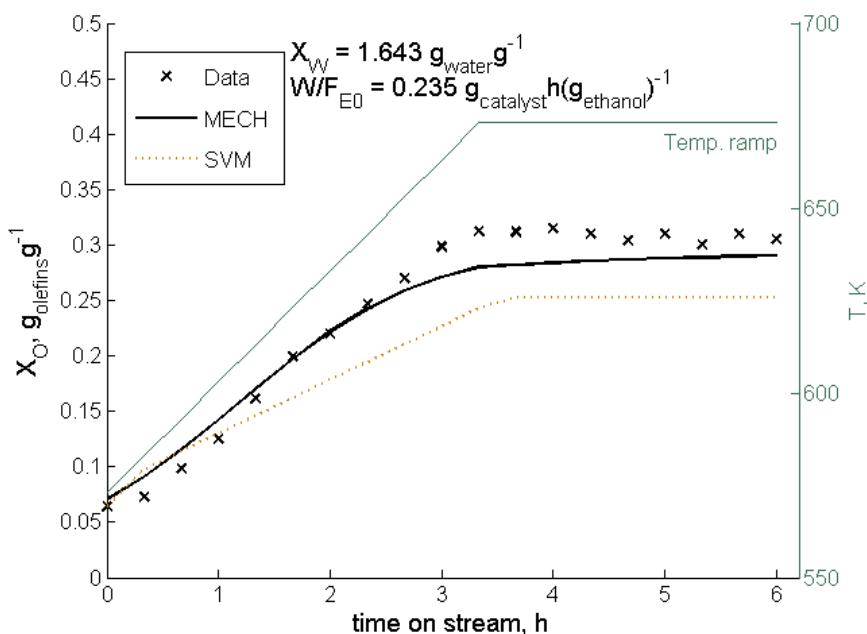


Figure 33. Example of the SVR based model performance for a variable temperature experiment

5.1.3. Artificial Neural Networks

As previously stated, based on the recommendations made by the literature regarding the modelling of chemical processes, various neural structures (see section 4.2.2.1 Structural models) have been implemented and iteratively improved. In this process, the sensitivity of each input on the model estimates has been analysed. Therefore, starting from simple networks, the neural model has been improved incorporating different variables in order to achieve reasonable values of the estimation error. Figure 34 shows the most representative implemented structural models. The starting point was the BASE structure, whose estimates are based only on the current values of the input variables. From this first structure, new structures have been created: DYN in which past values of these variables are included to provide the network with dynamic information, or REC network in which previously estimated output values are included as an inputs in a recurrent loop. TEMP further develops the NARX structure of

REC including the reaction time in order to incorporate the catalyst deactivation effect and its influence on the conversion of olefins. In this manner, information about the transformation rate is also provided to the model and its possible influence on the model estimates is checked.

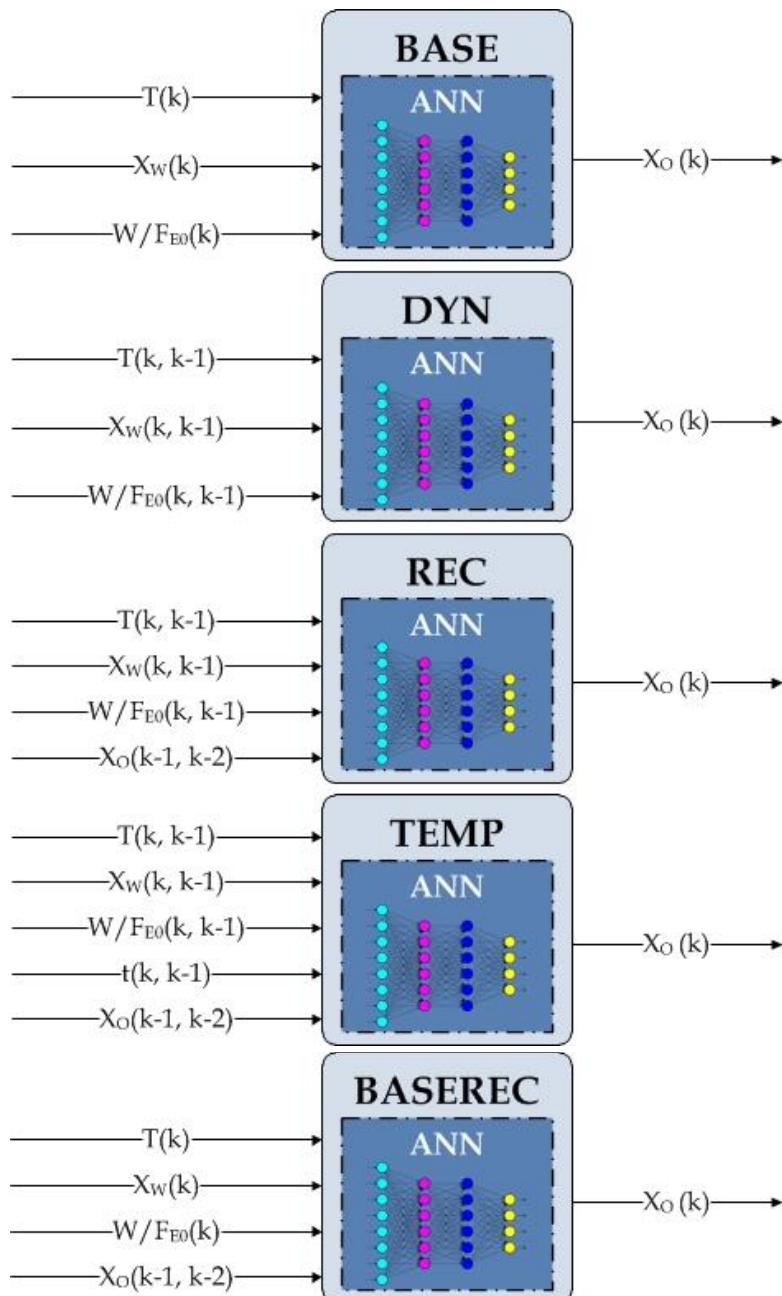


Figure 34. Most promising studied ANN configurations

Moreover, several variants of these structures have been tested with more or less memory of the process previous behaviour until the best combination of memory structure and dynamic information of the process has been set. More neural structures have also been tested with unsuccessful results, such as the BASEREC structure (see Figure 34). This network combines the input data of the BASE network with the recurrent loop with the previously estimated output values. However, this type of networks has provided poorer results than the finally selected structures.

For each structural model different internal structures have been tested, changing the number of hidden layers, the number of neurons of each layer and the output activation function. The size of the input and output layers come determined by the inputs of the model structure and the desired output of the model (X_0) respectively. The LOOCV procedure has been repeated several times and performed with 1 to 3 hidden layers, trying structures from 1 to 25 nodes in the hidden layers for each structure.

Additionally to modifying the neural structure, different activation functions have been tested. The objective has been to test if the use of non-linear activation functions improves the obtained results. Therefore three different output transfer functions have been tested (see Figure 35): the linear (*purelin*), the log-sigmoid non-linear function (*logsig*) and the hyperbolic tangent sigmoid function (*tansig*).

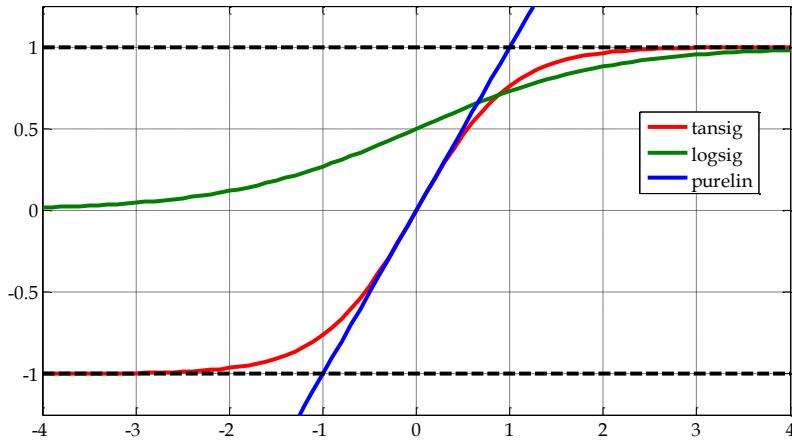


Figure 35. Transfer function during the training procedure

The best results have been found just with one hidden layer, the use of two or three hidden layers makes more complex the model without any improvement. Additionally, according to the literature, only one hidden layer is enough to approximate any continue function of n real variables (Cybenko 1989; Hornik et al. 1990).

Figure 36 to Figure 39 show the cross-validation average performance result with each structure. As can be seen, in near all the cases the best results are obtained with a reduced number of hidden neurons. In a first stage the average error decreases until the minimum value is obtained with the most suitable neural structure. Any other training procedure with a higher number of hidden neurons increases the cross-validation error measurement.

Table 8 shows the selected internal topology for each neural structure obtained with the LOOCV procedure. For all the structures the use of the hyperbolic tangent sigmoid function reduces the required number of hidden neurons. In some cases the best structure seems to be enough with just one hidden neuron. However, these structures present in all the cases the highest generalization error. While the structures trained with linear and log-sigmoid functions obtain similar results and significantly lower

generalization error comparing to the hyperbolic tangent sigmoid function. The best internal topology with linear transfer functions appears to be between 5 to 7 hidden neurons. The log-sigmoid function presents a larger range which goes from 5 hidden neurons with BASE and REC structures to 24 hidden neurons with the TEMP structure.

Table 8. Resume of the best internal topology for each structure and transfer function

| Structure | Linear | Log-sigmoid | Hyperbolic tangent sigmoid |
|-----------|--------|-------------|----------------------------|
| BASE | 3-7-1 | 3-5-1 | 3-4-1 |
| DYN | 6-5-1 | 6-10-1 | 6-3-1 |
| REC | 8-5-1 | 8-5-1 | 8-1-1 |
| TEMP | 10-6-1 | 10-24-1 | 10-1-1 |

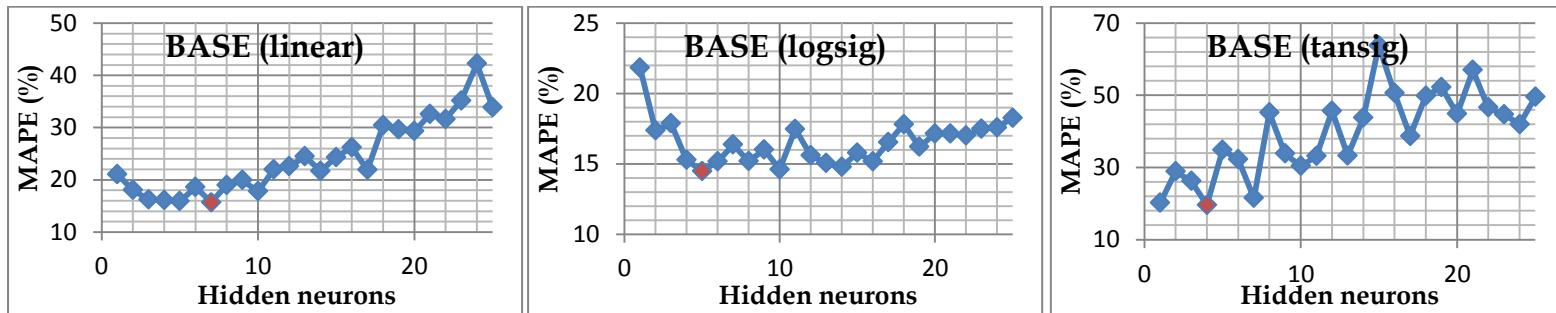


Figure 36. Neural structure selection with the LOOCV procedure for the BASE type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions.

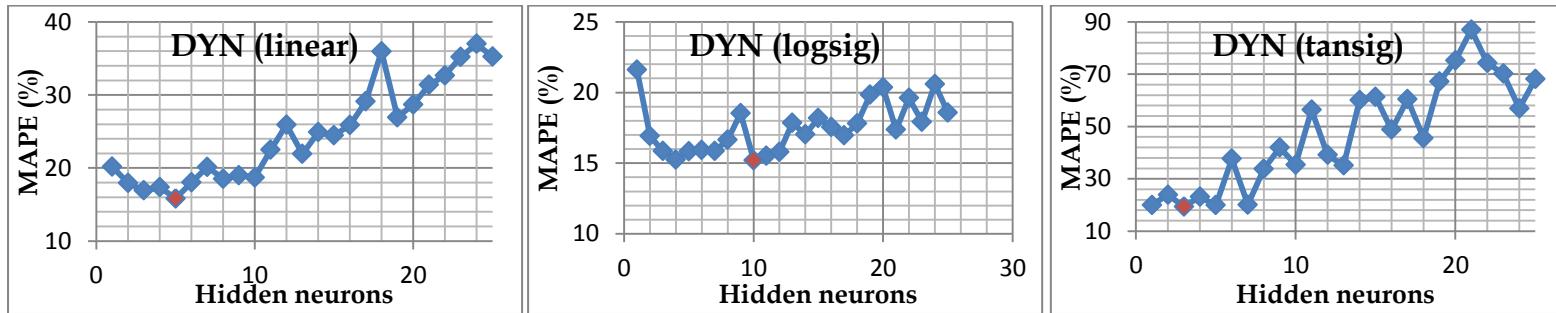


Figure 37. Neural structure selection with the LOOCV procedure for the DYN type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions.

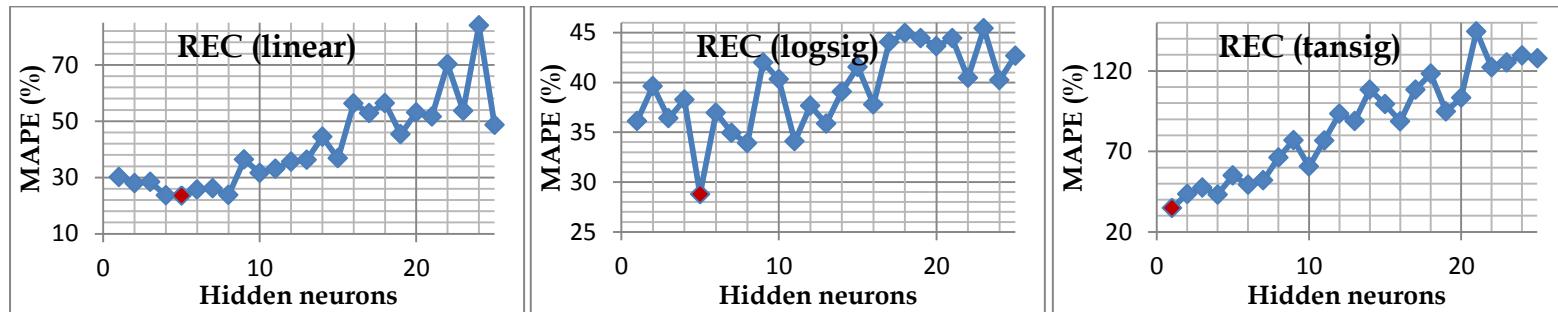


Figure 38. Neural structure selection with the LOOCV procedure for the REC type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions.

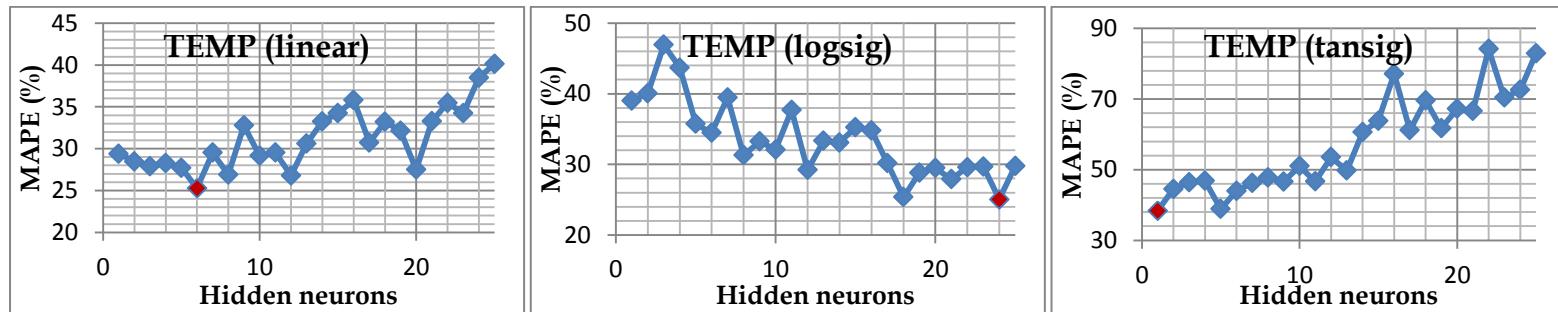


Figure 39. Neural structure selection with the LOOCV procedure for the TEMP type model with linear, hyperbolic tangent sigmoid or log-sigmoid output functions.

In addition to the selection of the ANN structure and internal topology, another very important aspect with a clear influence on the final results is the training functions. During the training process, the network weights are initialized, and then adjusted iteratively using the training functions to minimize the mean square error (MSE) between estimates and target values.

Table 9 shows the training algorithms used in the present work. Most of them use gradient-based techniques to determine how to adjust the network weights. The speed and performance of the training will vary depending on the used training algorithm.

Table 9. Training algorithms used during the neural training procedures

| Training algorithm |
|--|
| Bayesian regularization |
| Scaled conjugate gradient backpropagation |
| Random order incremental training with learning functions |
| Levenberg-Marquardt backpropagation |
| Cyclical order incremental training with learning functions |
| Polak-Ribiere conjugate gradient backpropagation |
| Gradient descent with momentum and adaptive lr backpropagation |
| Powell-Beale conjugate gradient backpropagation |
| Resilient backpropagation |
| Fletcher-Powell conjugate gradient backpropagation |
| BFGS quasi-Newton backpropagation |
| Gradient descent with adaptive lr backpropagation |
| One step secant backpropagation |
| Batch training with weight bias learning rules |
| Gradient descent with momentum backpropagation |
| Sequential order incremental training with learning functions |
| Gradient descent backpropagation |

Once selected the number of hidden neurons, the same LOOCV procedure has been used to select the most suitable training function. Table 10 to Table 13 show the LOOCV average results with each training function

for each model structure and used transfer function. Huge differences can be found between the algorithms, with average errors that range from 8,94 % (Bayesian Regularization) with the DYN structure to 137,32 % (Gradient descent with momentum backpropagation) in term of MAPE with TEMP type structure. It is significant that for both BASE and DYN model structures and independently to the used transfer function, the best training algorithm seems to be the Bayesian regularization. However, for the REC and TEMP structures the most suitable training function it is not so clear. In these two cases the cross-validation errors increase with all training functions and different algorithms seems to be the best option depending on the used transfer function.

Therefore, the best combination of model structure and training function results to be the BASE and DYN models with the Bayesian regularization with both linear or log-sigmoid output transfer function. Note that the differences in terms of RMSE and MAPE when comparing with different model structures and training algorithms are significant (Table 10 - Table 13) and in some cases have risen up to 143 %.

Table 10. Comparison of the LOOCV average results obtained for each training algorithm for BASE type model structure.

| Training algorithm | Linear | | Log-sigmoid | | hyperbolic tan-sigmoid | |
|--|----------|--------------------------|-------------|--------------------------|------------------------|--------------------------|
| | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) |
| Bayesian regularization | 10,349 | 0,019888 | 10,505 | 0,020156 | 11,506 | 0,021314 |
| Scaled conjugate gradient backpropagation | 34,738 | 0,054112 | 26,173 | 0,040897 | 31,348 | 0,047186 |
| Random order incremental with learning functions | 22,058 | 0,03395 | 28,019 | 0,039227 | 23,979 | 0,036434 |
| Levenberg-Marquardt backpropagation | 18,306 | 0,031882 | 15,636 | 0,026104 | 28,127 | 0,04354 |
| Cyclical order incremental with learning functions | 44,953 | 0,078693 | 38,587 | 0,05615 | 34,814 | 0,054216 |
| Polak-Ribiere conjugate gradient backpropagation | 28,272 | 0,045878 | 21,26 | 0,034457 | 35,194 | 0,05246 |
| Gradient descent with momentum and adaptive lr backpropagation | 105,68 | 0,1572 | 32,793 | 0,043769 | 74,963 | 0,10428 |
| Powell-Beale conjugate gradient backpropagation | 28,629 | 0,044756 | 20,062 | 0,032704 | 51,034 | 0,074984 |
| Resilient backpropagation | 49,92 | 0,076936 | 22,772 | 0,032368 | 36,38 | 0,055433 |
| Fletcher-Powell conjugate gradient backpropagation | 29,361 | 0,04925 | 21,018 | 0,032228 | 43,864 | 0,065939 |
| BFGS quasi-Newton backpropagation | 27,206 | 0,044068 | 24,992 | 0,037251 | 33,012 | 0,050941 |
| Gradient descent with adaptive lr backpropagation | 70,154 | 0,10354 | 28,421 | 0,03747 | 62,108 | 0,080276 |
| One step secant backpropagation | 30,533 | 0,049488 | 26,94 | 0,041578 | 58,083 | 0,080291 |
| Batch training with weight bias learning rules | 118,92 | 0,17881 | 64,861 | 0,078549 | 74,559 | 0,11235 |
| Gradient descent with momentum backpropagation | 128,92 | 0,18392 | 68,001 | 0,080525 | 68,847 | 0,1001 |
| Sequential order incremental with learning functions | 50,101 | 0,071519 | 67,274 | 0,084429 | 45,827 | 0,065651 |
| Gradient descent backpropagation | 88,383 | 0,13413 | 67,259 | 0,081274 | 61,046 | 0,093882 |

Table 11. Comparison of the LOOCV average results obtained for each training algorithm for DYN type model structure

| Training algorithm | Linear | | Log-sigmoid | | hyperbolic tan-sigmoid | |
|--|----------|--------------------------------|-------------|--------------------------------|------------------------|--------------------------------|
| | MAPE (%) | RMSE (golefins ⁻¹) | MAPE (%) | RMSE (golefins ⁻¹) | MAPE (%) | RMSE (golefins ⁻¹) |
| Bayesian regularization | 9,9511 | 0,019544 | 8,9438 | 0,018097 | 11,754 | 0,022509 |
| Scaled conjugate gradient backpropagation | 26,734 | 0,042707 | 25,578 | 0,044931 | 30,003 | 0,046826 |
| Random order incremental with learning functions | 22,272 | 0,034732 | 22,435 | 0,03397 | 21,953 | 0,034131 |
| Levenberg-Marquardt backpropagation | 17,047 | 0,031562 | 16,702 | 0,029119 | 29,225 | 0,043393 |
| Cyclical order incremental with learning functions | 37,64 | 0,06208 | 33,852 | 0,052252 | 29,195 | 0,046226 |
| Polak-Ribiere conjugate gradient backpropagation | 25,353 | 0,040134 | 19,876 | 0,034261 | 27,396 | 0,041731 |
| Gradient descent with momentum and adaptive lr backpropagation | 71,73 | 0,10782 | 29,976 | 0,043055 | 45,423 | 0,069542 |
| Powell-Beale conjugate gradient backpropagation | 23,324 | 0,037327 | 20,712 | 0,034911 | 30,391 | 0,04479 |
| Resilient backpropagation | 32,183 | 0,059454 | 20,724 | 0,0345 | 22,637 | 0,037259 |
| Fletcher-Powell conjugate gradient backpropagation | 25,034 | 0,041994 | 23,137 | 0,03755 | 38,81 | 0,064386 |
| BFGS quasi-Newton backpropagation | 19,266 | 0,03196 | 26,563 | 0,043707 | 24,826 | 0,038888 |
| Gradient descent with adaptive lr backpropagation | 58,292 | 0,085236 | 27,47 | 0,041945 | 41,087 | 0,063201 |
| One step secant backpropagation | 26,095 | 0,041236 | 28,819 | 0,046532 | 22,359 | 0,03569 |
| Batch training with weigh bias learning rules | 74,387 | 0,10997 | 55,007 | 0,076998 | 50,845 | 0,077443 |
| Gradient descent with momentum backpropagation | 85,448 | 0,1235 | 55,618 | 0,077475 | 56,889 | 0,08082 |
| Sequential order incremental with learning functions | 39,533 | 0,059817 | 56,477 | 0,081564 | 44,254 | 0,061409 |
| Gradient descent backpropagation | 59,131 | 0,09302 | 59,666 | 0,081076 | 56,56 | 0,082066 |

Table 12. Comparison of the LOOCV average results obtained for each training algorithm for REC type model structure

| Training algorithm | Linear | | Log-sigmoid | | hyperbolic tan-sigmoid | |
|--|---------------|--------------------------|---------------|--------------------------|------------------------|--------------------------|
| | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) |
| Bayesian regularization | 42,349 | 0,058441 | 44,033 | 0,06235 | 36,361 | 0,045501 |
| Scaled conjugate gradient backpropagation | 37,682 | 0,056172 | 26,743 | 0,041744 | 23,463 | 0,035696 |
| Random order incremental with learning functions | 24,471 | 0,037403 | 25,657 | 0,036966 | 26,407 | 0,038766 |
| Levenberg-Marquardt backpropagation | 45,931 | 0,07218 | 33,588 | 0,05129 | 32,215 | 0,043608 |
| Cyclical order incremental with learning functions | 41,495 | 0,067092 | 31,901 | 0,047319 | 28,825 | 0,042584 |
| Polak-Ribiere conjugate gradient backpropagation | 32,13 | 0,05284 | 20,622 | 0,030722 | 23,755 | 0,035971 |
| Gradient descent with momentum and adaptive lr backpropagation | 92,055 | 0,14987 | 28,591 | 0,039121 | 38,851 | 0,057614 |
| Powell-Beale conjugate gradient backpropagation | 31,996 | 0,050861 | 19,307 | 0,029577 | 20,873 | 0,031187 |
| Resilient backpropagation | 61,319 | 0,099348 | 20,848 | 0,031509 | 25,829 | 0,037916 |
| Fletcher-Powell conjugate gradient backpropagation | 30,668 | 0,047914 | 21,408 | 0,033449 | 22,76 | 0,033592 |
| BFGS quasi-Newton backpropagation | 28,332 | 0,043119 | 22,797 | 0,034702 | 20,078 | 0,029963 |
| Gradient descent with adaptive lr backpropagation | 80,908 | 0,1308 | 29,667 | 0,040477 | 33,737 | 0,050656 |
| One step secant backpropagation | 32,253 | 0,04868 | 26,249 | 0,039088 | 22,231 | 0,033841 |
| Batch training with weight bias learning rules | 114,76 | 0,16847 | 63,988 | 0,077219 | 48,06 | 0,068966 |
| Gradient descent with momentum backpropagation | 102,1 | 0,16284 | 58,42 | 0,069788 | 37,064 | 0,052354 |
| Sequential order incremental with learning functions | 50,616 | 0,081346 | 55,087 | 0,071625 | 42,125 | 0,056062 |
| Gradient descent backpropagation | 75,668 | 0,11554 | 56,325 | 0,071087 | 39,631 | 0,055878 |

Table 13. Comparison of the LOOCV average results obtained for each training algorithm for TEMP type model structure

| Training algorithm | Linear | | Log-sigmoid | | hyperbolic tan-sigmoid | |
|--|--------------|--------------------------|---------------|--------------------------|------------------------|--------------------------|
| | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) | MAPE (%) | RMSE (golefins $^{-1}$) |
| Bayesian regularization | 35,674 | 0,052224 | 6,5999 | 0,015938 | 37,848 | 0,045954 |
| Scaled conjugate gradient backpropagation | 29,415 | 0,050033 | 27,381 | 0,048602 | 32,371 | 0,040683 |
| Random order incremental with learning functions | 24,669 | 0,037813 | 22,493 | 0,034756 | 25,54 | 0,038558 |
| Levenberg-Marquardt backpropagation | 43,904 | 0,061803 | 29,241 | 0,051021 | 38,069 | 0,046948 |
| Cyclical order incremental with learning functions | 39,168 | 0,06673 | 39,11 | 0,061476 | 29,637 | 0,043222 |
| Polak-Ribiere conjugate gradient backpropagation | 30,242 | 0,051411 | 27,782 | 0,051133 | 23,916 | 0,034399 |
| Gradient descent with momentum and adaptive lr backpropagation | 152,01 | 0,23491 | 43,203 | 0,075523 | 45,901 | 0,064103 |
| Powell-Beale conjugate gradient backpropagation | 31,317 | 0,051621 | 27,672 | 0,048347 | 23,444 | 0,03381 |
| Resilient backpropagation | 71,783 | 0,12198 | 22,727 | 0,042395 | 26,469 | 0,042808 |
| Fletcher-Powell conjugate gradient backpropagation | 35,141 | 0,061432 | 32,477 | 0,061029 | 24,528 | 0,03468 |
| BFGS quasi-Newton backpropagation | 24,25 | 0,038997 | 27,655 | 0,052131 | 22,067 | 0,032718 |
| Gradient descent with adaptive lr backpropagation | 86,511 | 0,14977 | 39,333 | 0,070244 | 34,343 | 0,05175 |
| One step secant backpropagation | 30,005 | 0,052634 | 34,073 | 0,061251 | 24,085 | 0,034337 |
| Batch training with weight bias learning rules | 131,01 | 0,22282 | 60,687 | 0,091203 | 40,557 | 0,065049 |
| Gradient descent with momentum backpropagation | 137,32 | 0,23439 | 54,876 | 0,091297 | 43,779 | 0,063618 |
| Sequential order incremental with learning functions | 59,756 | 0,090059 | 56,561 | 0,094796 | 42,584 | 0,058448 |
| Gradient descent backpropagation | 106,08 | 0,16869 | 59,731 | 0,099589 | 36,401 | 0,05713 |

Figure 40 shows the model architecture that seems to present the best results. Table 14 shows the resume of the RMSE and MAPE LOOCV average values for all experimental sets when being represented by the model structures in Figure 34 with the selected structure and training function. The results for linear and log-sigmoid output functions are only presented, being this one the best option. The used of the hyperbolic tangent sigmoid function as output transfer function has been discarded due to its poor results. Note that the optimum internal topology for each structure is also indicated.

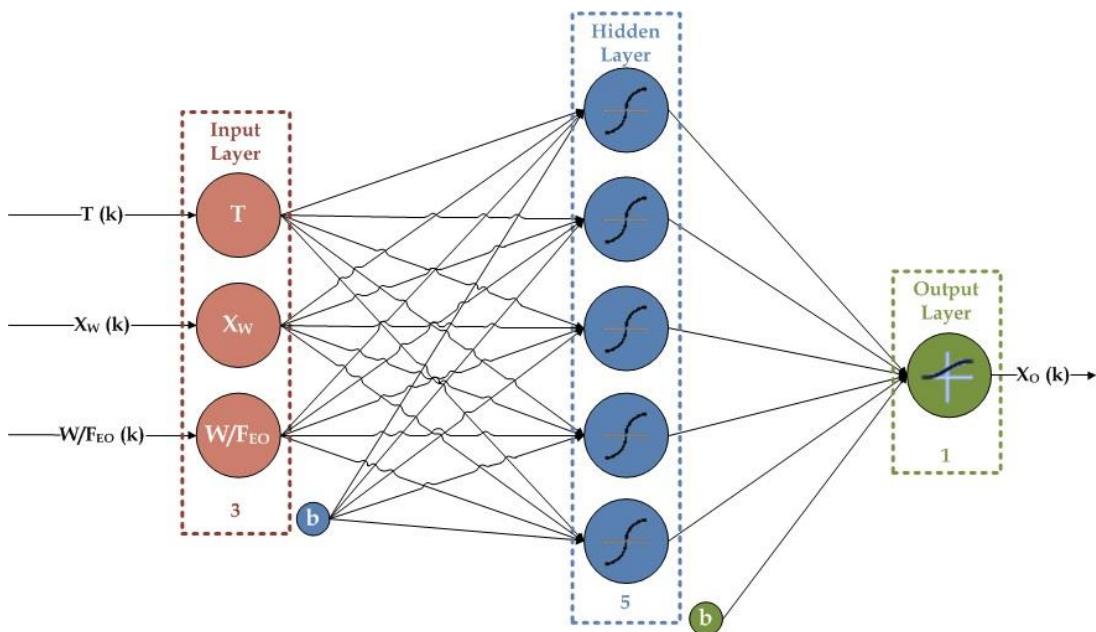


Figure 40. Neural model architecture (BASE) pre-selected for the BTO process

Based on these initial results, the BASE model with 7 nodes in the hidden layer (3-7-1) and the DYN model with 10 nodes in the hidden layer (6-10-1) present the best possible alternatives in terms of RMSE and MAPE average values when following a cross validation methodology. Both solutions present very close average errors for both lineal and log-sigmoid activation functions. BASE and DYN represent neural model structures that use as input variables the main variables affecting the conversion of olefins. The result are feedforward type ANN, using non-linear activation functions and

with a dynamic structure dependent on the current and past values of the input variables in the case of DYN model, and a more simple network depending just on the current values of the input variables (BASE model). The other two alternatives, REC and TEMP structures, have been discarded due to their poor results with all tested combinations of hidden neurons, output transfer function type and training algorithm.

Table 14. Resume of cross-validation estimation average errors for the main implemented models

| Output function | ANN structure (internal topology) | RMSE ($\text{golefins} \text{g}^{-1}$) | MAPE (%) |
|-----------------|-----------------------------------|--|----------|
| Linear | BASE (3-7-1) | 0.0198 | 10.34 |
| | DYN (6-5-1) | 0.0195 | 9.95 |
| | REC (8-5-1) | 0.0374 | 24.47 |
| | TEMP (10-6-1) | 0.0389 | 24.25 |
| Log-sigmoid | BASE (3-5-1) | 0.0201 | 10.50 |
| | DYN (6-10-1) | 0.0180 | 8.94 |
| | REC (8-5-1) | 0.0295 | 19.30 |
| | TEMP (10-24-1) | 0.0347 | 22.49 |

Table 15 shows the results obtained by the BASE and DYN models, and those obtained with the MECH mechanistic model when estimating the production of olefins for experiments carried out at T_{const} and T_{ramp} as detailed in Table 2. The neural models present better results than those obtained by the mechanistic model both in terms of RMSE, MAE and MAPE when using both T_{const} and T_{ramp} conditions. In addition, it uses much lower estimation time which will allow its use for the optimization and control of industrial applications of olefin production. Indeed, the BASE and DYN neural models reduce the estimation times up to 65 times compared to the mechanistic model.

Table 15. Estimation mean errors obtained for all experiments with MECH, BASE and DYN models

| Models | | RMSE (golefins g^{-1}) | MAE (golefins g^{-1}) | MAPE (%) | HITS (%) |
|--------|---------------------------------------|------------------------------|-----------------------------|----------------|----------------|
| MECH | T _{const} | 0.0313 | 0.0248 | 13.9927 | 53.7117 |
| | T _{ramp} | 0.0337 | 0.0285 | 18.3752 | 36.7088 |
| | T _{const} +T _{ramp} | 0.0323 | 0.0263 | 15.7886 | 46.6494 |
| BASE | T _{const} | 0.0146 | 0.0122 | 7.0798 | 81.8518 |
| | T _{ramp} | 0.0293 | 0.0249 | 16.3678 | 52.7319 |
| | T _{const} +T _{ramp} | 0.0204 | 0.0172 | 10.7142 | 70.4570 |
| DYN | T _{const} | 0.0134 | 0.0113 | 6.5122 | 83.5839 |
| | T _{ramp} | 0.0253 | 0.0202 | 12.7224 | 64.2817 |
| | T _{const} +T _{ramp} | 0.0181 | 0.0147 | 8.9423 | 76.0309 |

Finally, in order to validate the fitted neural models and to test its generalization capability, several completely new operational conditions have been simulated with the mechanistic knowledge model to prepare an additional test dataset. Therefore, the neural models with the designed structure, topology and training functions are trained with the complete training dataset used in the LOOCV procedure and validated with the new and unseen test dataset. The fitted models show test MAPE and RMSE values of 11.88 % and 0.0338 golefins g^{-1} respectively for the BASE structure and 15.11 % and 0.0425 golefins g^{-1} for DYN neural structure. Table 16 shows that the DYN model reaches better performance than the BASE model for the experimental training data. However, the BASE model with the log-sigmoid output activation function presents the best results for the test dataset and for the global performance. It seems that the BASE model with linear output activation function is not flexible enough to outperform DYN model structures. However, the use of a non-linear output activation function provides to the BASE model the necessary dynamism to improve the results by only adding a little extra complexity to the model; while pushing the DYN model to overfit the training data. Although being dynamic, the BTO process is a relatively slow process. Thus, the BASE neural networks have been able to outperform dynamic models within the experimental

operational range. Figure 41 shows the experimental and model results, which shows a good level of agreement with a coefficient of determination (R^2) of 0.8526 for the BASE model shown in Figure 40. The conclusion is that the soft-modelling (BASE) estimates both the real experimental data and the test simulated data with the desired accuracy and faster than the mathematical model. These results are similar to those obtained by other related works. To the author's knowledge, there are not any other research on the same process, but in similar processes very promising results were obtained following the same neural modelling procedure for other kind of chemical reactors (Kashani & Shahhosseini 2010). Feedforward neural networks have been successfully used to model other catalytic processes and the catalyst performance (Holeňa & Baerns 2003). However, the studied research contributions related to the BTO process are based on mechanistic models based on reaction kinetics. The developed neural model outperforms this mechanistic model which presents a 15.7886 % of MAPE error and a R^2 of 0.8055 when comparing with the experimental data (Gayubo, Alonso, Valle, Andrés T. Aguayo, et al. 2010; Gayubo et al. 2012). In addition, the literature presents several chemical and biological complex processes that have been modelled successfully using neural models with similar neural structures (Dach et al. 2016; López et al. 2017; Mikulandrić et al. 2014) and with similar performance (Nair et al. 2016).

Table 16. Validation average errors with BASE and DYN models for Train, Test and Total datasets

| Output function | ANN structure | MAPE (%) | | R^2 |
|-----------------|---------------|----------|--------|--------|
| | | Train | Test | |
| Linear | BASE | 10.271 | 17.562 | 0.7626 |
| | DYN | 9.452 | 12.410 | 0.8517 |
| Log-sigmoid | BASE | 10.714 | 11.887 | 0.8526 |
| | DYN | 8.942 | 15.110 | 0.8341 |

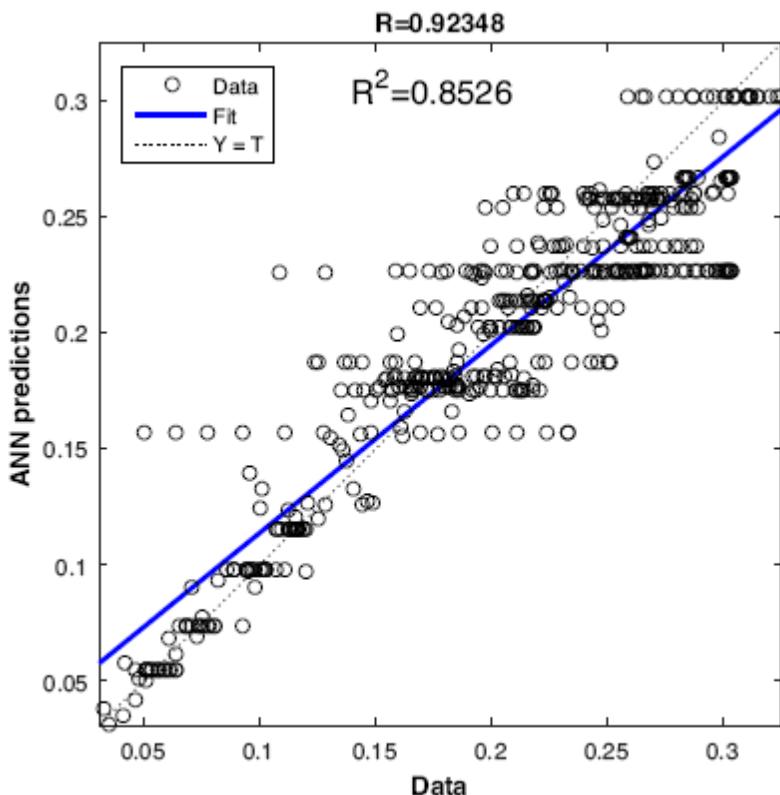


Figure 41. Regression curve of the model BASE (3-5-1) for the entire dataset

Figure 42 and Figure 43 show, as an example, the simulated and experimental results of four experiments (two using temperature ramps experiments and two with constant temperature). It can be clearly observed that the mechanistic (MECH) and neural network (BASE and DYN) models show a similar and satisfactory level of agreement. Thus, the ANN model has proved to be a powerful model structure in view of assimilating and generalizing the system dynamics.

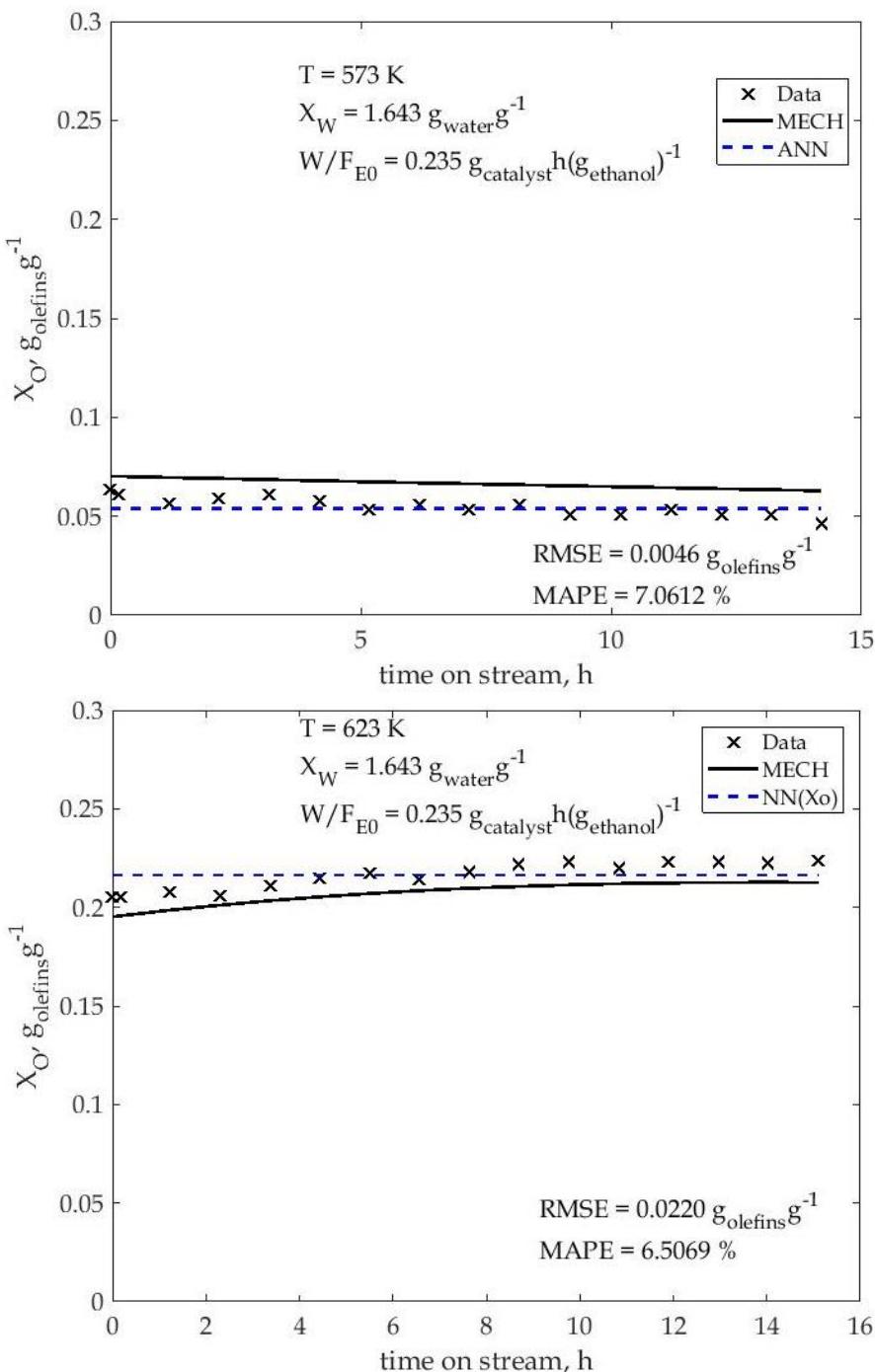


Figure 42. Comparison between the simulation and experimental results of olefins concentration when using constant temperature conditions. The simulation results correspond to a mechanistic model (MECH) and neural networks based model (ANN).

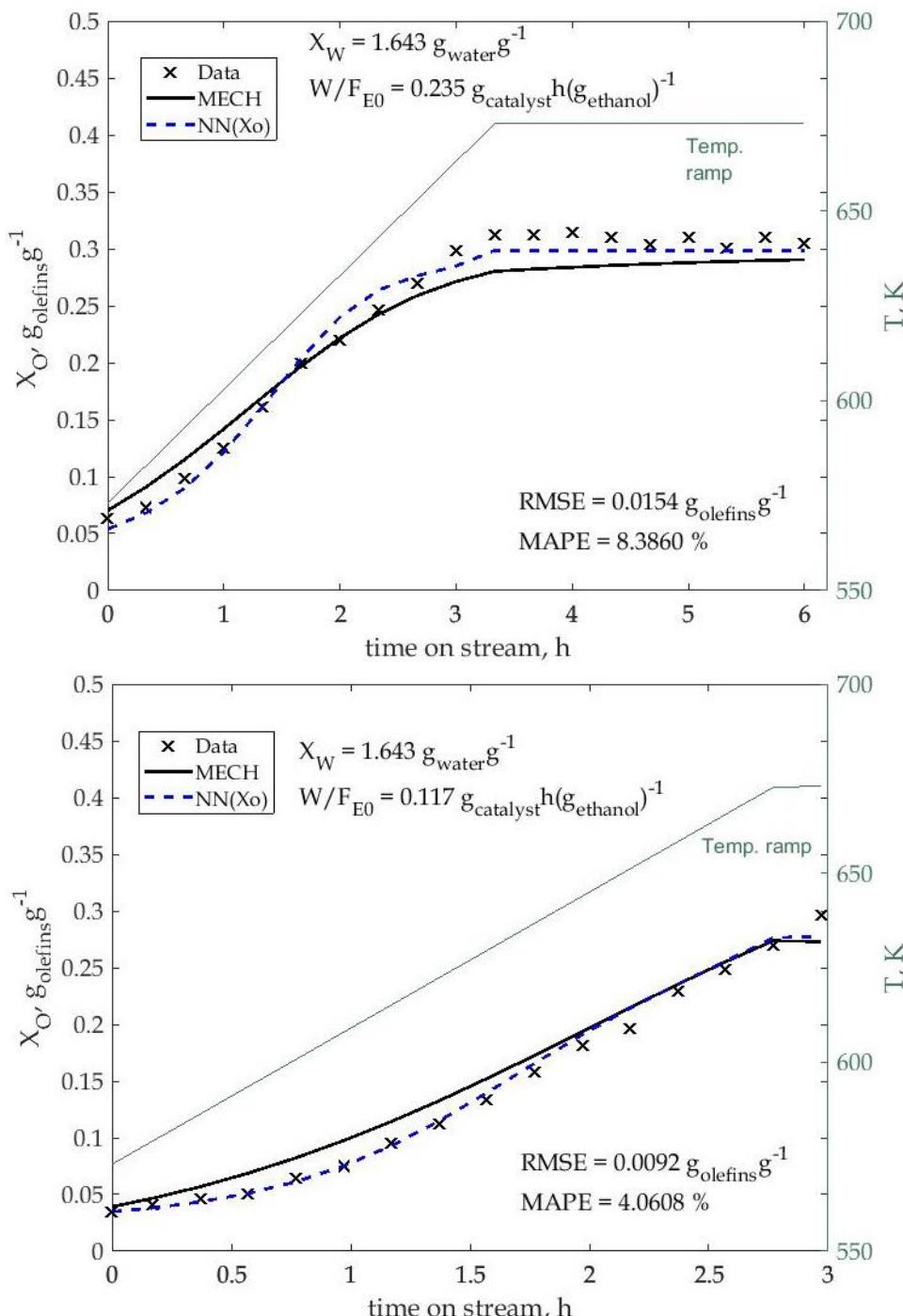


Figure 43. Comparison between the simulation and experimental results of olefins concentration when using temperature ramps. The simulation results correspond to a mechanistic model (MECH), a neural networks based model (ANN).

5.1.4. Hybrid Neural Model

A key point to understand the BTO process is the evolution of the activity of the catalyst used to promote the process. The catalyst deactivation rate is a non-directly measurable effect of the accumulation of coke by the catalyst, blocking their active centres. Therefore is not possible to obtain experimental data in order to be used in the modelling procedure to incorporate information about the deactivation. However, due to the relevance of the deactivation rate increase with the production time, it can be interesting to introduce this information in the model (especially for long-term estimations).

In order to provide the model with the information of both kinetics of the process (the main reaction and the deactivation kinetic), a hybrid modelling approach has been selected. The hybrid models, also known as grey-box models, combine black-box type models with traditional knowledge-based mathematical models. Usually the most complex or unknown parts of the system is modelled using the selected soft-computing technique while for the well-known parts knowledge models are used.

Therefore, Hybrid Neural Networks models (HNN) combine elements of knowledge models with ANN to model the unknown and most complex parts of a process (Molga 2003).

Figure 44 shows the conceptual scheme of the proposed hybrid model. It is a combination of a well-known part of the knowledge model to estimate the kinetic of the catalyst deactivation, with a neural network based model to assimilate the kinetic of the main reaction taking into account the information of the catalyst deactivation. As the previous models, the neural network will receive information from the main process variables and, depending on the neural structure type, previously generated estimations (see Figure 45). Additionally, the neural network receives as new input the activity rate in each time step, generated by the kinetic catalyst deactivation

knowledge model. The model of the deactivation kinetic estimates the activity value based on the temperature and the water content in the feed.

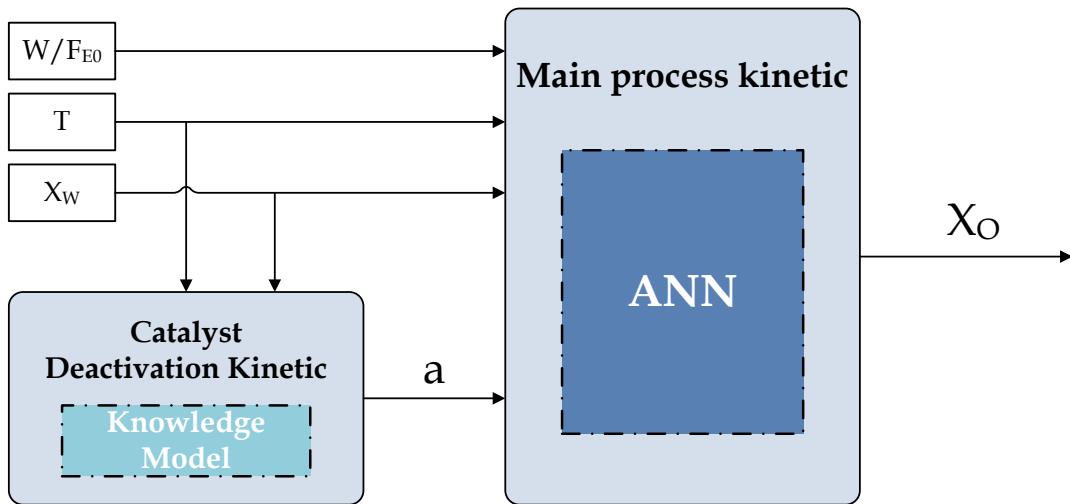


Figure 44. Conceptual scheme of the hybrid modelling approach with ANN.

In addition, for this new hybrid modelling approach the synthetic dataset, which increase the dynamic information about the BTO process provided to the model, will be used during the training procedure (see subchapter 4.1.3). These new data will provide to the trained models the necessary dynamic information about the behaviour of the process in long-term experiments.

Using the new dataset, in order to select the most suitable neural structure for the new hybrid neural model, the previously explained LOOCV procedure has been used. Figure 45 shows two model structures proposed for the HNN model. These structures are based on the BASE and REC type models, taking into account the previously obtained results (see subchapter 5.1.1). The model MOD 1 is based on the REC type structure with the main process variables (T , X_W and W/F_{E0}) as inputs, the activity (a) and the previous values of the olefins conversion rate (X_O). For the activity the current and several previous values are provided to the neural network. The model MOD 2 is based on the BASE structure including data about the current and past activity values. Finally, as in the previous modelling

procedure, both structures have been tested using two output activation functions; linear and log-sigmoid functions. The hyperbolic tangent sigmoid activation function has been discarded due to the poor results obtained in the previous modelling procedure.

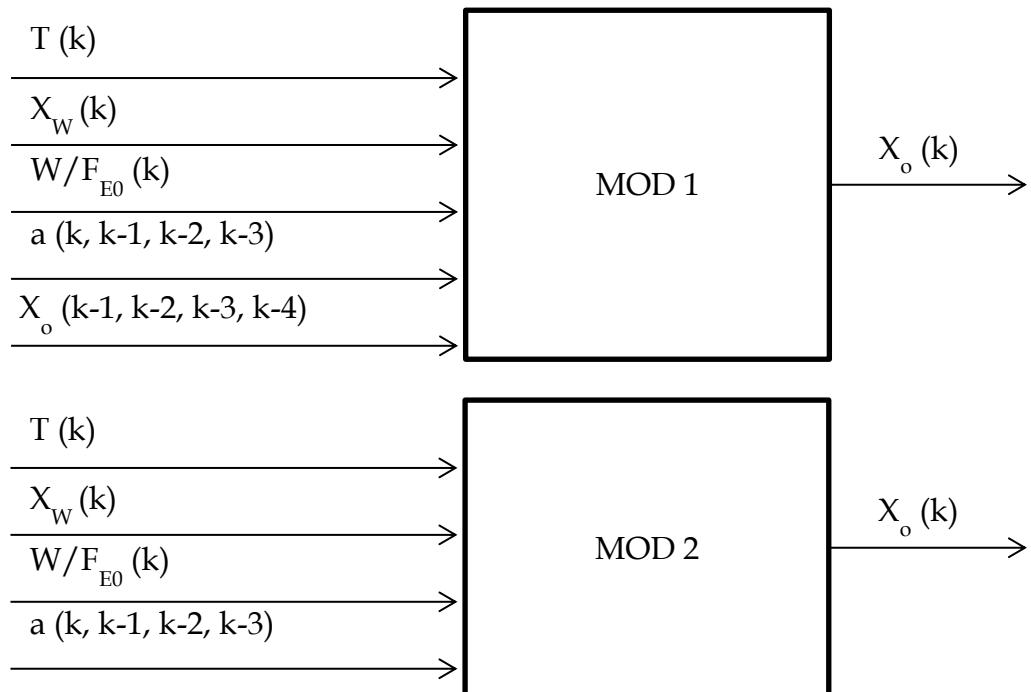


Figure 45. Model structures proposed for the hybrid modelling approach.

Figure 46 and Figure 47 shows the LOOCV average MAPE errors when applying the LOOCV procedure for MOD 1 and MOD 2 with both linear and non-linear log-sigmoid output activation functions and with one hidden layer with from 1 to 25 neurons. All of them reach to a minimum error value before reaching to 25 hidden neurons. In the case of the MOD 1, with both linear and log-sigmoid activation function, the best results are obtained with relatively high neurons; 22 and 23 hidden neurons respectively. In contrast,

MOD 2 presents the best results with 11 and 20 hidden neurons with the linear and log-sigmoid activation function respectively.

Once selected the best neural structure, the most suitable training function has been selected. From the previously tested training algorithms (see Table 9) the most promising ones have been selected. Those which have obtained the best results in the first modelling procedure:

- Bayesian Regularization
- Levenberg-Marquardt backpropagation
- Polak-Ribiere conjugate gradient backpropagation
- Powell-Beale conjugate gradient backpropagation
- Fletcher-Powell conjugate gradient backpropagation

Table 17 shows the results of the LOOCV procedure carried out with each of the selected training algorithms. In all cases, for both MOD 1 and MOD 2 models, and with both linear and log-sigmoid activation functions, the best results are obtained by the Bayesian Regularization algorithm. In the comparison of the output activation function, the best performance measurements are obtained with the log-sigmoid functions. They obtain a mean cross-validation MAPE errors of 4.36 % and 3.8 % for the MOD 1 and MOD 2 models respectively. In contrast, using the linear function the results obtained by the proposed models are slightly worse but still present good performance. The MOD 2 presents slightly better results than MOD 1 even than not having the feedback loop of olefins conversion rate (X_O).

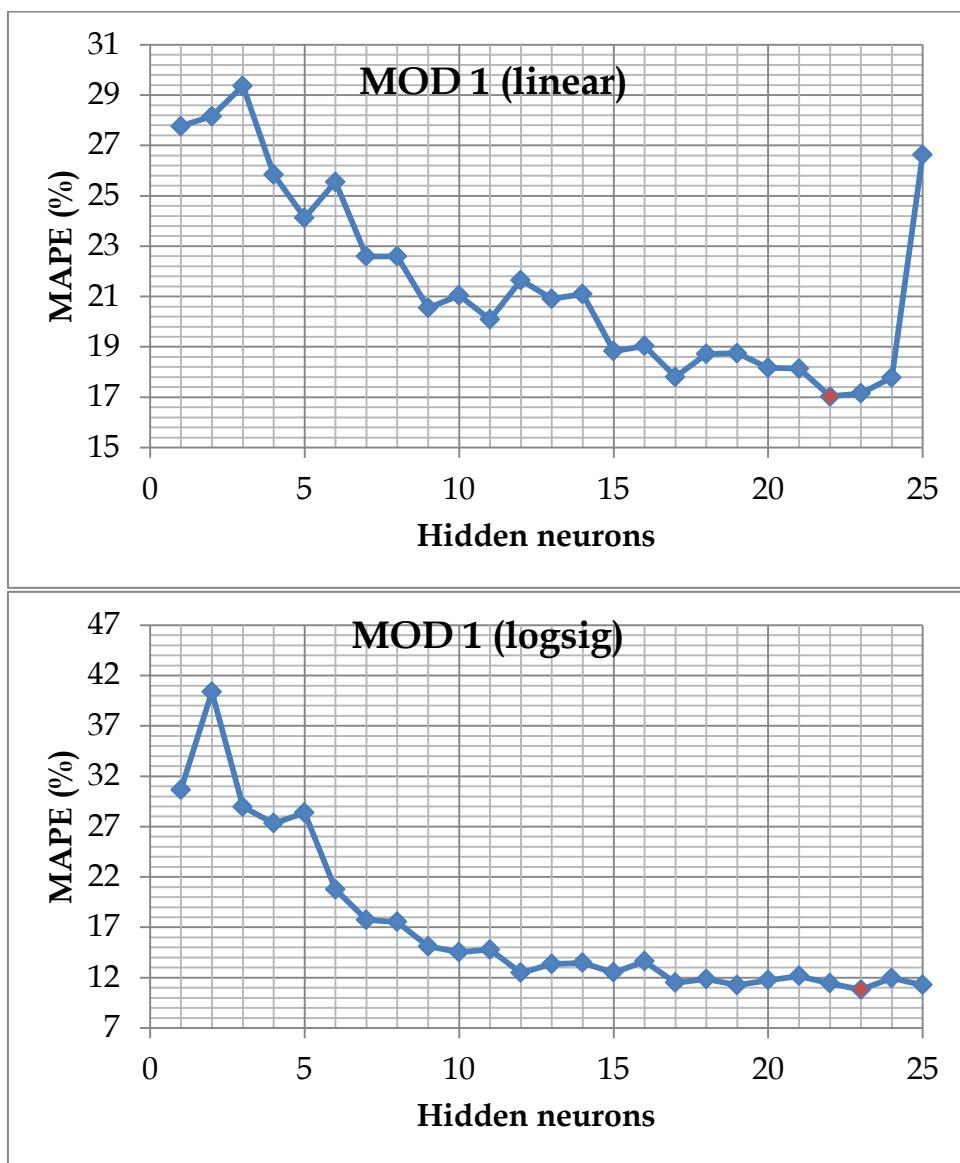


Figure 46. Neural structure selection with the LOOCV procedure for the MOD 1 type model with linear and log-sigmoid output transfer functions.

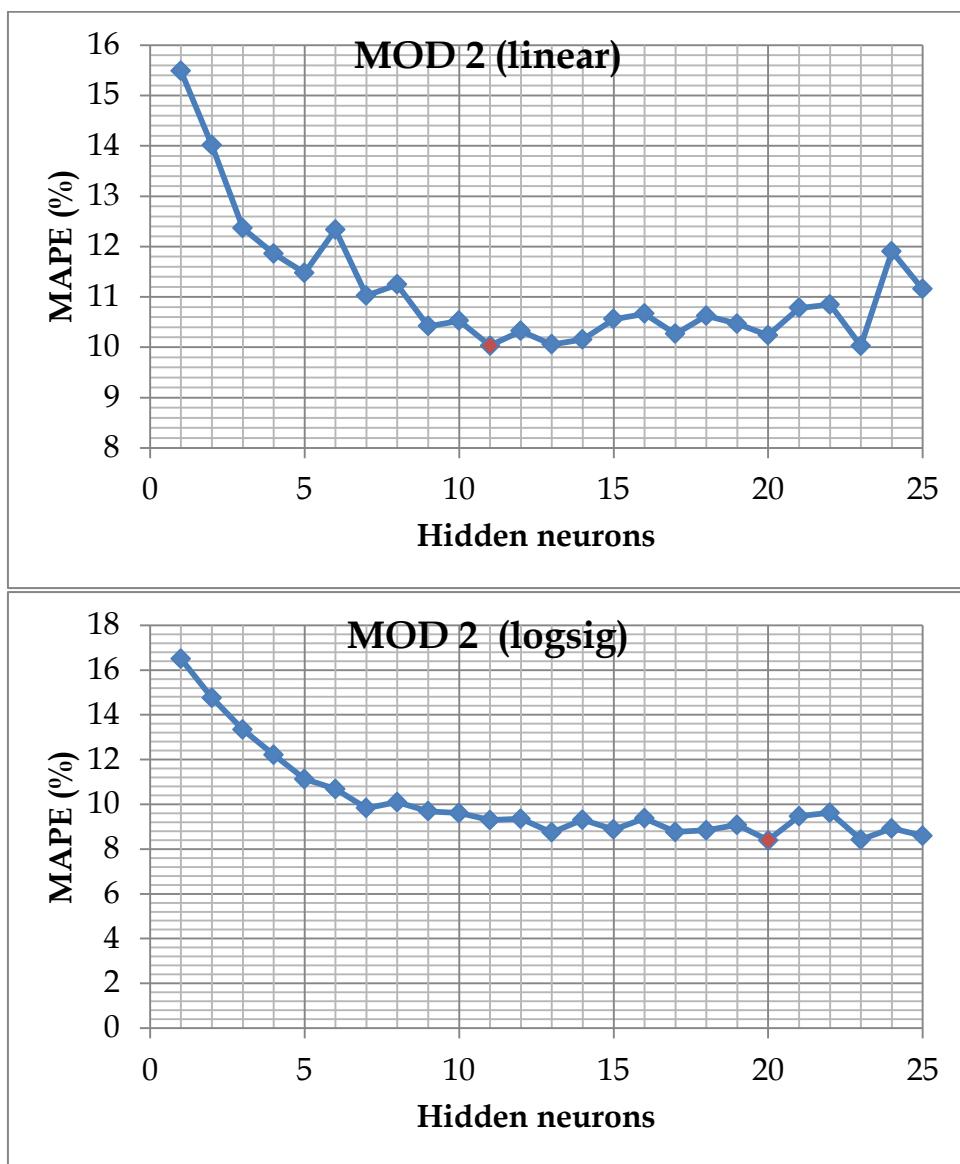


Figure 47. Neural structure selection with the LOOCV procedure for the MOD 2 type model with linear and log-sigmoid output transfer functions.

Table 17. Comparison of the LOOCV average results obtained for the most suitable training algorithms for both MOD 1 and MOD 2 models.

| Model | Training algorithm | Linear | | Log-sigmoid | |
|-------|--|--------------|---------------------------------|--------------|---------------------------------|
| | | MAPE (%) | RMSE (golefins^{-1}) | MAPE (%) | RMSE (golefins^{-1}) |
| MOD 1 | Bayesian Regularization | 5,515 | 0,010289 | 4,361 | 0,008724 |
| | Levenberg-Marquardt backpropagation | 16,951 | 0,029511 | 16,381 | 0,030301 |
| | Polak-Ribiere conjugate gradient backpropagation | 275,931 | 0,520031 | 16,470 | 0,030122 |
| | Powell-Beale conjugate gradient backpropagation | 212,397 | 0,461302 | 17,689 | 0,034887 |
| | Fletcher-Powell conjugate gradient backpropagation | 239,218 | 0,421972 | 19,957 | 0,037931 |
| MOD 2 | Bayesian Regularization | 4,563 | 0,009021 | 3,801 | 0,007937 |
| | Levenberg-Marquardt backpropagation | 9,095 | 0,017765 | 14,645 | 0,028174 |
| | Polak-Ribiere conjugate gradient backpropagation | 21,806 | 0,041201 | 17,423 | 0,032201 |
| | Powell-Beale conjugate gradient backpropagation | 17,656 | 0,034350 | 15,285 | 0,028684 |
| | Fletcher-Powell conjugate gradient backpropagation | 26,257 | 0,044918 | 21,867 | 0,041327 |

Table 18 shows the resume of the RMSE and MAPE LOOCV average values for all training data sets when being represented by the model structures in Figure 45 with the selected structure and training function. From these results, the log-sigmoid function has been selected as the most suitable output activation function. The MOD 2 models seems to obtain the best solution, with an internal topology of 7-20-1, the log-sigmoid output activation function and using the Bayesian regularization as the training algorithm.

Table 18. Resume of the Cross-Validation results for the selected structures.

| Output function | HNN structure (internal topology) | RMSE (golefins $^{-1}$) | MAPE (%) |
|-----------------|--------------------------------------|-----------------------------|----------|
| Linear | MOD 1 (11-22-1) | 0,010289 | 5,515 |
| | MOD 2 (7-11-1) | 0,009021 | 4,563 |
| Log-sigmoid | MOD 1 (11-23-1) | 0,008724 | 4,361 |
| | MOD 2 (7-20-1) | 0,007937 | 3,801 |

Once the neural structures and their hyperparameters, such as the activation functions or training functions, have been selected, the next step has been the model validation. The selected hybrid neural structures have been trained with all the available data except the test dataset. These experimental conditions have been maintained unseen for the models in the whole training and structure selection procedure.

Table 19 shows the goodness-of-fit of the selected models. Both models present similar training and test results. The training MAPE average errors are 0.70 % and 0.37 % for MOD 1 and MOD 2 models respectively. As test MAPE errors, MOD 1 and MOD 2 show a 3.26 % and 3.63 % average error respectively. In terms of RMSE for the test dataset, show an average error of 0.0088 golefins $^{-1}$ and 0.0066 golefins $^{-1}$ respectively. These absolute errors are lower than the experimental error, which was determined as 0.02 golefins $^{-1}$ (see section 4.2). Therefore, it can be conclude that both model structures are suitable and perform very similarly. Finally, in Figure 48 and Figure 49, and Figure 50 and Figure 51 the regression curves of both models for the training

and test datasets are presented. In all cases, the models show a good level of agreement with a coefficient of determination (R^2) always over 0.98. They obtain a 0.9991 and 0.9856 with MOD 1 model for training and test data respectively and 0.9998 and 0.9918 respectively for MOD 2 model.

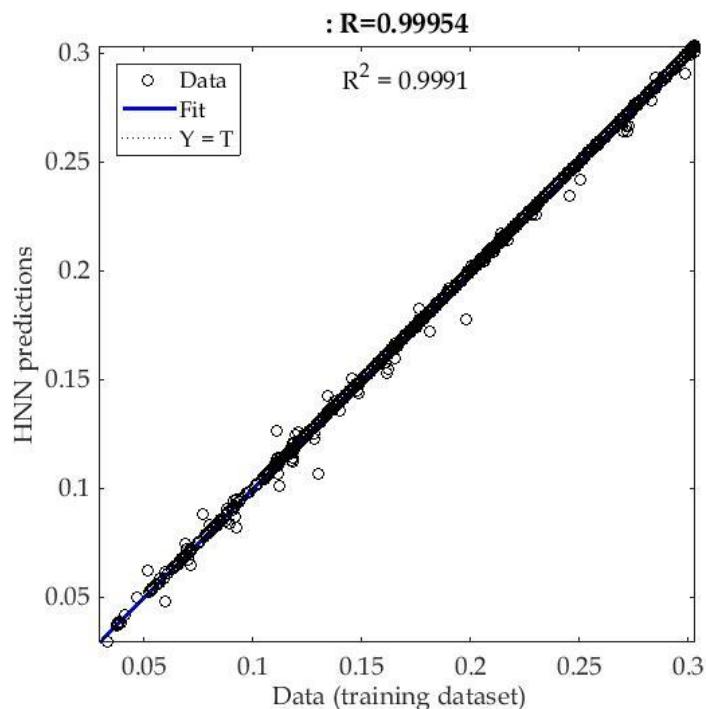


Figure 48. Regression curve of the model MOD 1 (11-23-1) for the training dataset.

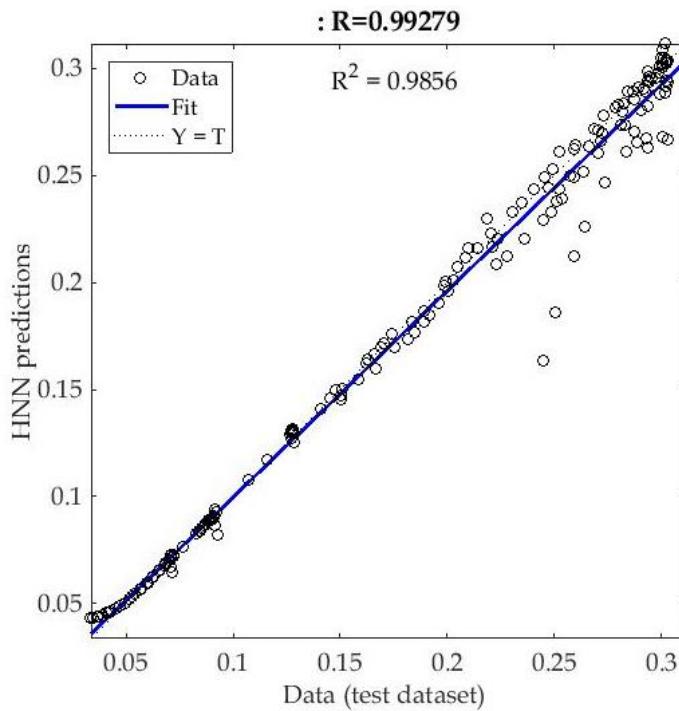


Figure 49. Regression curve of the model MOD 1 (11-23-1) for the test dataset.

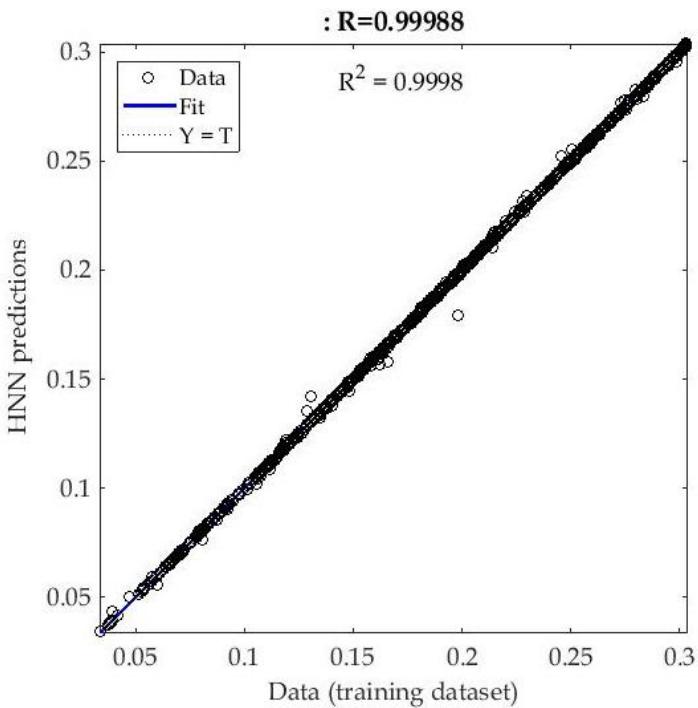


Figure 50. Regression curve of the model MOD 2 (7-20-1) for the training dataset.

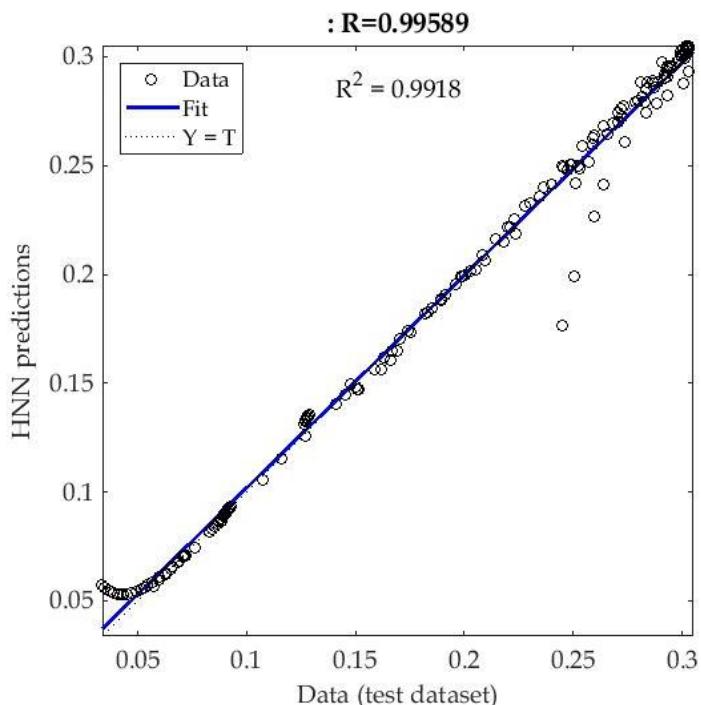


Figure 51. Regression curve of the model MOD 2 (7-20-1) for the test dataset.

Table 19. Goodness-of-fit of the models for the Test dataset.

| Output function | HNN structure | MAPE (%) | | R^2 | |
|-----------------|---------------|----------|--------|----------|--------|
| | | Training | Test | Training | Test |
| Log-sigmoid | MOD 1 | 0.7074 | 3.2692 | 0.9991 | 0.9856 |
| | MOD 2 | 0.3732 | 3.6343 | 0.9998 | 0.9918 |

The current training and validation datasets, combine the original short-term experimental conditions with new data of the process behaviour that prolong the necessary forecasting horizons of the models of those presented in the original experimental data (see section 4.1.2). Despite to this fact that hinders the modelling objectives, the obtained models are able to show a satisfactory level of agreement. Additionally, the results obtained for the original experimental conditions are even better than any previous result.

Table 20 shows the goodness-of-fit of the developed hybrid models for the original experimental conditions simulated by the mechanistic models. The average errors in terms of MAPE and RMSE are 0.93% and 0.002223 g_{olefins}g⁻¹ respectively for the MOD 1 model; and 0.48 % and 0.001155 g_{olefins}g⁻¹ respectively for the MOD 2 model.

Table 20. Goodness-of-fit of the hybrid models to the original operational conditions.

| Output function | HNN structure (internal topology) | RMSE (g _{olefins} g ⁻¹) | MAPE (%) |
|-----------------|-----------------------------------|--|----------|
| Log-sigmoid | MOD 1 (11-23-1) | 0.002223 | 0.9372 |
| | MOD 2 (7-20-1) | 0.001155 | 0.4831 |

The level of agreement with the original operational conditions is impressively closed. Figure 52 and Figure 53 show the regression curves of MOD 1 and MOD 2 models for those operational conditions. They present a coefficient of determination of 0.9983 and 0.9994 respectively.

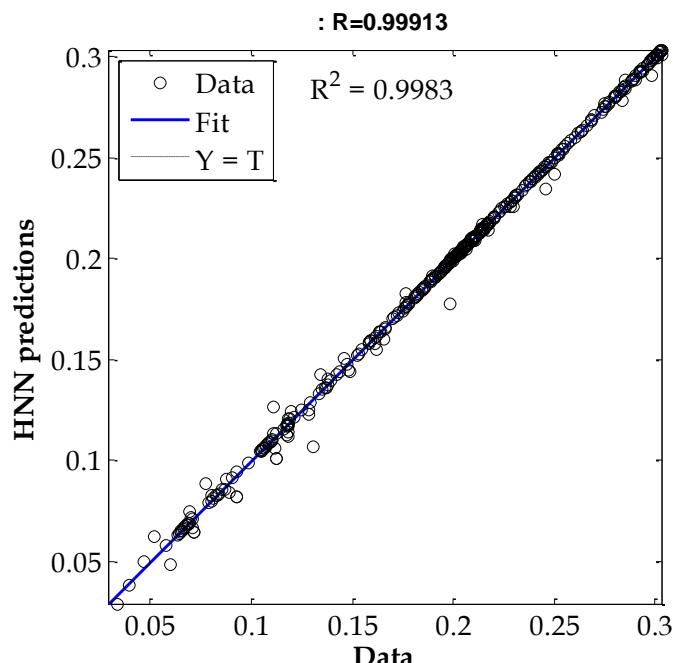


Figure 52. Regression curve of the model MOD 1 (11-23-1) for the simulated data of the original operational conditions

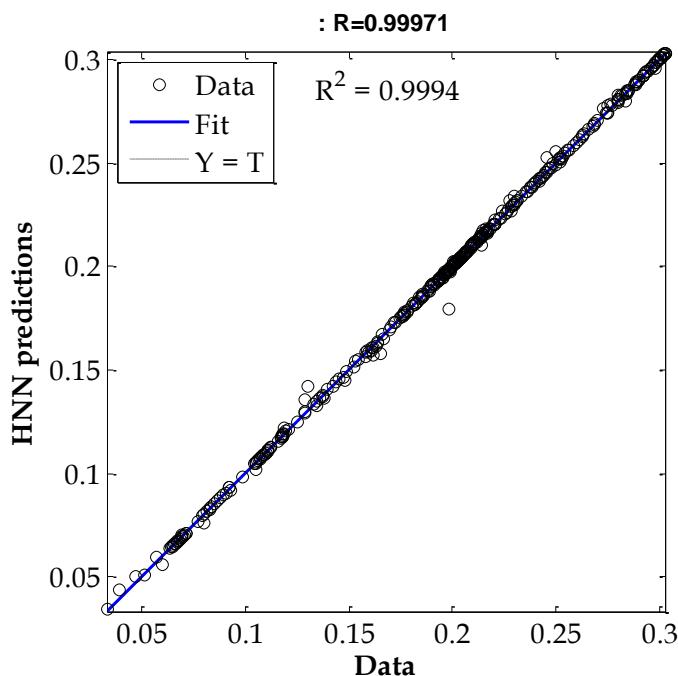


Figure 53. Regression curve of the model MOD 2 (7-20-1) for the simulated data of the original operational conditions

Since the training and validation procedure has been performed over the synthetic experimental data (see section 4.1.3), an additional analysis has been carried out to test if the fitted models are able to reproduce the real behaviour of the BTO process with the same level of agreement than the mechanistic model. Table 21 shows the goodness-of-fit of the hybrid neural models and of the mechanistic model to the original experimental data. All of them present similar average error measurements, with MAPE average error of 15.7886 %, 15.8684 % and 15.9145 % for the MECH, MOD 1 and MOD 2 models respectively. The estimations made by the HNN models over the real data present only 0.0798 (MOD 1) and 0.1259 (MOD 2) percentage points higher errors than the MECH model.

Table 21. Model comparison for all the available real experimental data.

| Models | RMSE (golefins g^{-1}) | MAE (golefins g^{-1}) | MAPE (%) | HITS (%) |
|---------------|-------------------------------------|------------------------------------|--------------------|--------------------|
| MECH | 0.0323 | 0.0263 | 15.7886 | 46.6494 |
| MOD1 | 0.0324 | 0.0264 | 15.8684 | 44.5876 |
| MOD2 | 0.0325 | 0.0265 | 15.9145 | 45.8762 |

Finally, as an example, Figure 54 presents the estimations of the MOD 1 and MOD 2 neural model, which have obtained similar results, for two different sets of experiments with complex trajectories of the operating variables. On the top panel continue temperature variations are tested and on the bottom panel both temperature and the water content on the feed are changed simultaneously. These results show the capacity of the HNN models to properly assimilate and reproduce the BTO process dynamics even with simultaneous changes in the process variables.

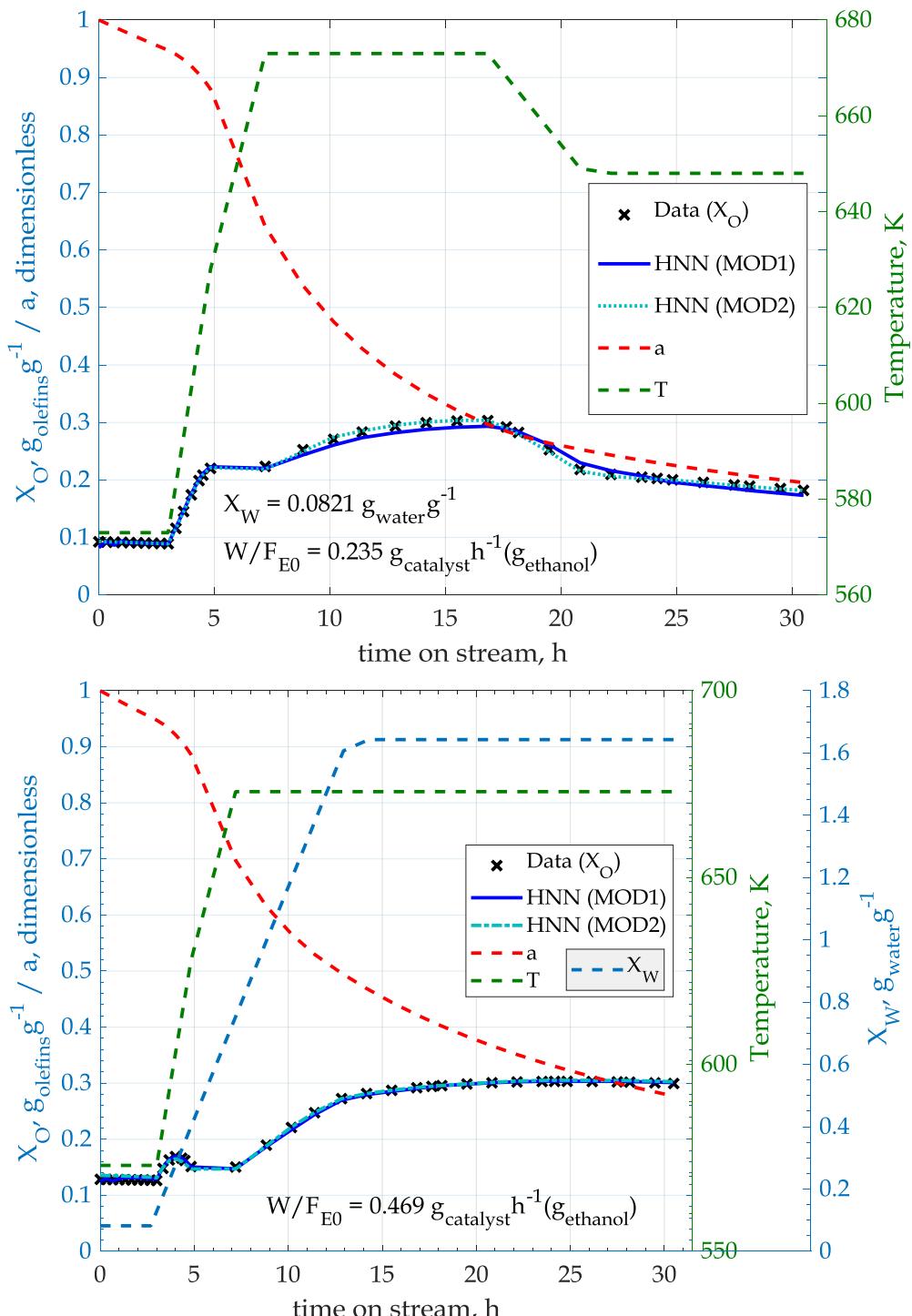


Figure 54. HNN model estimations for the evolution of the BTO process for an experimental run with complex trajectories for the variables T (top panel) and T and X_W (bottom panel)

5.2. Model Comparison

In order to compare the developed models, the behaviour of the BTO process under several new operational conditions has been simulated using the MECH model. These new test dataset has been used to compare all the previously developed models under the same operational conditions for both description and prediction of the dynamic behaviour of the BTO process. This new test dataset includes complex and extended experiments, where several operating variables are change simultaneously during the simulation. The objective is to test the behaviour of the models under these types of complex operational conditions that will appear in reality (see section 4.2.2.2).

Some models such as the first ANN based models, have shown very good results to describe the process behaviour under the constant experimental operational conditions. They have been properly validated and tested, and their generalization capability analysed within that initial experimental operational conditions. However, the optimisation procedure requires estimates with higher prediction horizons and with varying temperature profiles where these models are unable to provide acceptable estimates. The previous models are unable to correctly estimate the behaviour of the process when the catalyst lifespan is prolonged. These models need additional information, such as the activity data provided to the Hybrid based models, to be able to assimilate the real process dynamics.

Figure 55 shows the model design procedure carried out during the modelling of the BTO process (see section 4.2.2.2). Two critical aspects are analysed, the ability of each model to describe the experimental data and for long-term predictions of the BTO process dynamic behaviour. As previously presented, LR and SVM based models are not able to properly represent the behaviour of the process. Whereas, the ANN obtain acceptable results to describe the experimental data, they cannot properly predict the behaviour of the process when the catalyst is being deactivated. This aspect is critical

for the optimisation step, because the obtained model must be able to correctly simulate the process behaviour during the complete catalyst lifespan.

On the other hand, the hybrid modelling approach based on ANN, the HNN models, have been able to successfully predict the process behaviour in both situations.

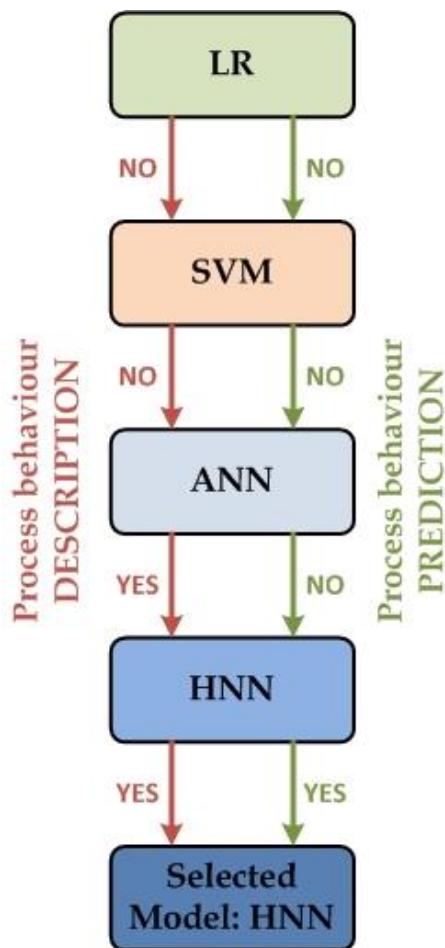


Figure 55. Prediction capability of the trained models

This fact occurs because the optimisation objective is to extend the catalyst lifespan and to maximize the total production. In order to achieve these objectives it is necessary to prolong the simulation time of the process

considerably higher. In this situation, the influence of the deactivation kinetic which previously was not so critical, takes relevance and becomes one of the principal drives of the process.

Therefore, the first optimisation results were not satisfactory because the error committed by the surrogate models become significant when the catalyst deactivation was prolonged. Additionally, the longer the production step is, the higher the effect of the catalyst deactivation over the total production of olefins, and the base initial models (LR, SVM and ANN) did not take into account the catalyst activity rate.

Table 22 shows the performance comparison of the most relevant models for the new test dataset for model comparison (see section 4.2.2.2). Those models, such as the hybrid models MOD 1 and MOD 2, which use the information about the catalyst lifespan to predict the behaviour of the BTO process present significantly better results than any other model proposal. On the one hand, MOD 1 and MOD 2 models, with 3.26 % and 3.63 % of MAPE average error respectively, overcome the performance of the ANN based model in more than 15 percentage points. On the other hand, ANN based model, with 18.73 % of MAPE error presents better results than those models based on LR and SVR techniques.

Table 22. Goodness-of-fit comparison for the best trained solution with each modelling technique for the same unseen operational conditions.

| Model | TEST dataset | |
|-----------|------------------------------|-------------|
| | RMSE (golefins g^{-1}) | MAPE (%) |
| LR | 0.041145 | 22.6240 |
| SVR | 0.050961 | 25.0852 |
| ANN | 0.044700 | 18.7394 |
| HNN MOD 1 | 0.008801 | 3.2692 |
| HNN MOD 2 | 0.006634 | 3.6343 |

Figure 56 and Figure 57 show, as an example, the behaviour of the process under different operational conditions included in the new test dataset. The important changes of each operating variable produce regimen

changes in the process dynamics that each model should detect. As can be seen, only the hybrid models are able to follow the behaviour of the process with the desired accuracy.

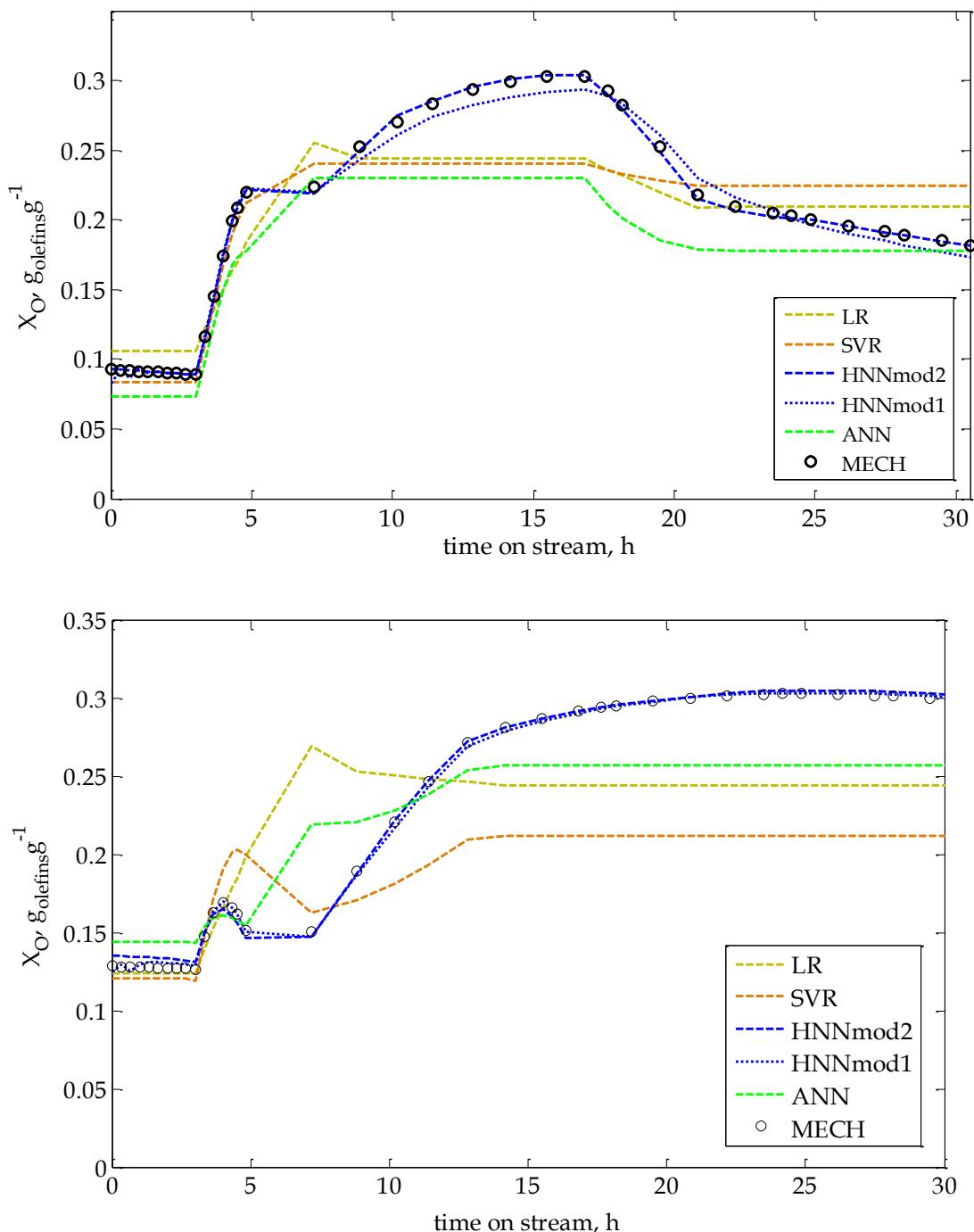


Figure 56. Comparison of the presented model's performance under new complex and unseen operational conditions.

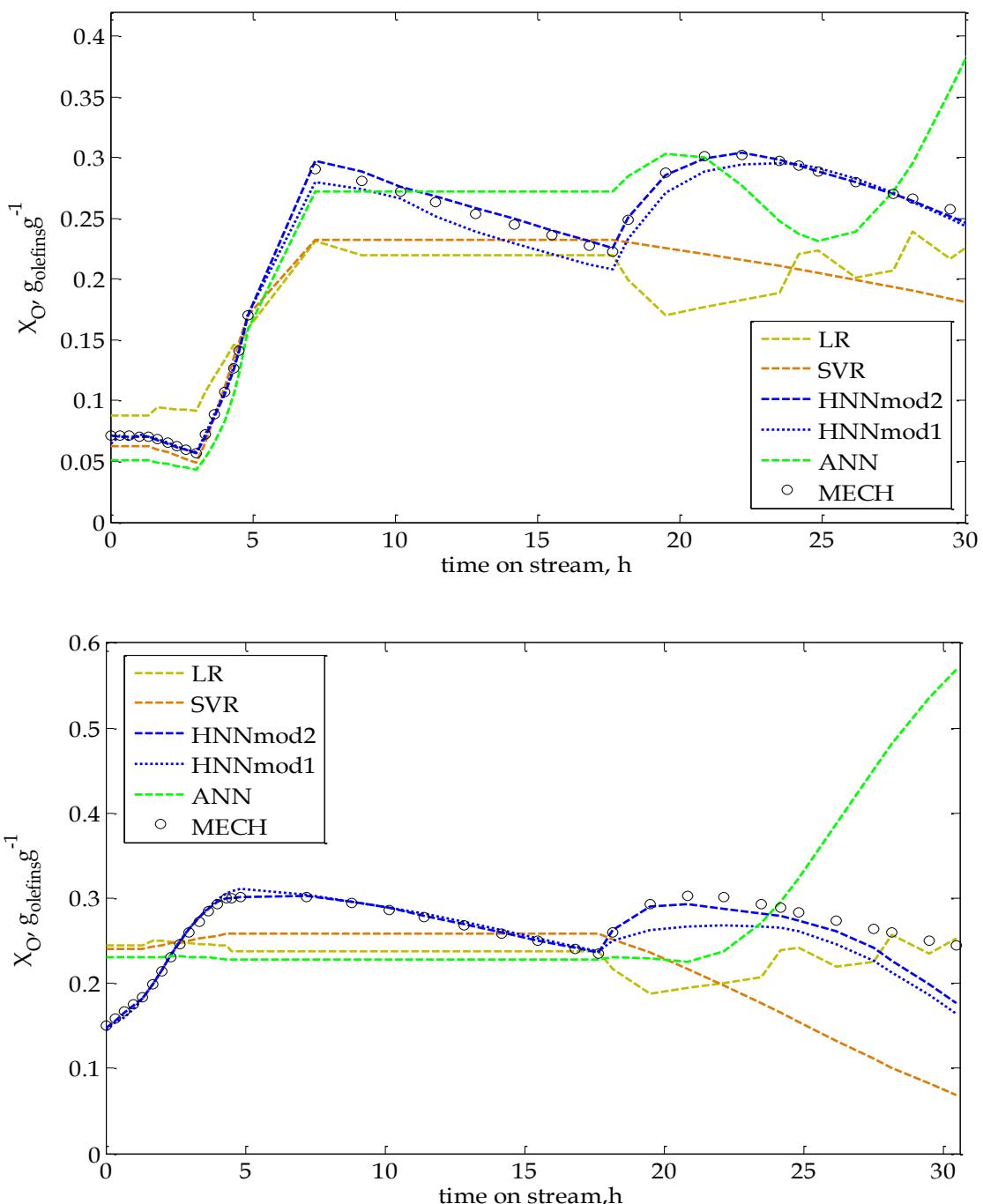


Figure 57. Comparison of the presented model's performance under new complex and unseen operational conditions.

Finally, Figure 58 and Figure 59 show the behaviour of each model under new unseen operational conditions with different prediction horizons. Different experiments with 20, 70, 200 and 450 hours as prediction horizons have been simulated. In all of them, the hybrid models are the only ones able to estimate the dynamics of the BTO process. Specifically, the MOD 2 hybrid model, which presents the best performance measurements, seems to be the best option in all the cases.

This model (HNN MOD 2) seems to be most promising solution. It is able to successfully estimate the complex operational conditions where several operating variables change simultaneously; in some cases with contrary effects on production. Moreover, it is able to predict the process behaviour with the desired accuracy for long-term estimations of more than 400 hours, when the catalyst deactivation rate is the main driver of the process. The model has been able to correctly identify the dynamics of the BTO process from the training data, and that is why it is able to predict with the desired accuracy both complex and long-term operational conditions.

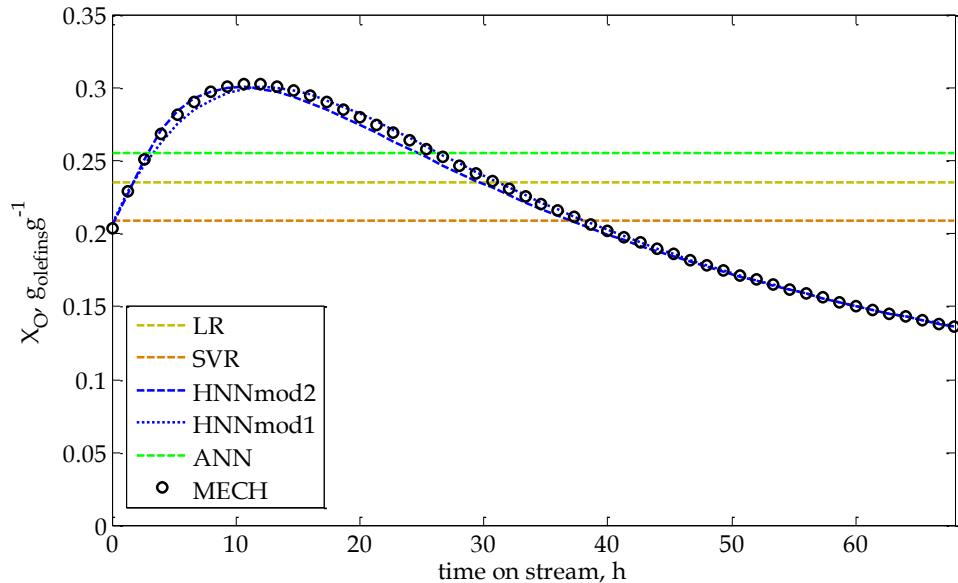
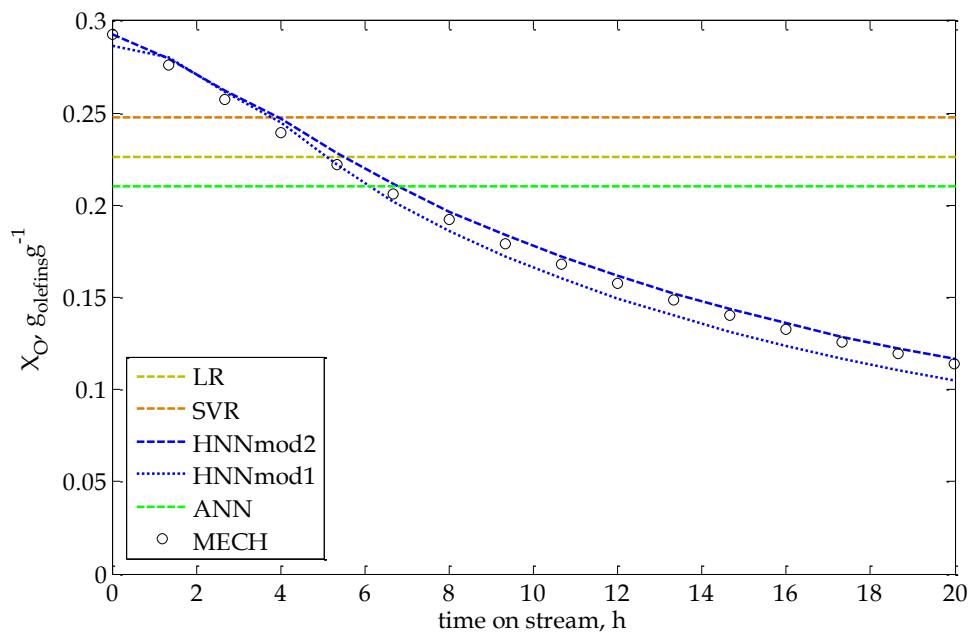


Figure 58. Comparison of the presented model's performance under new operational conditions with different prediction horizons.

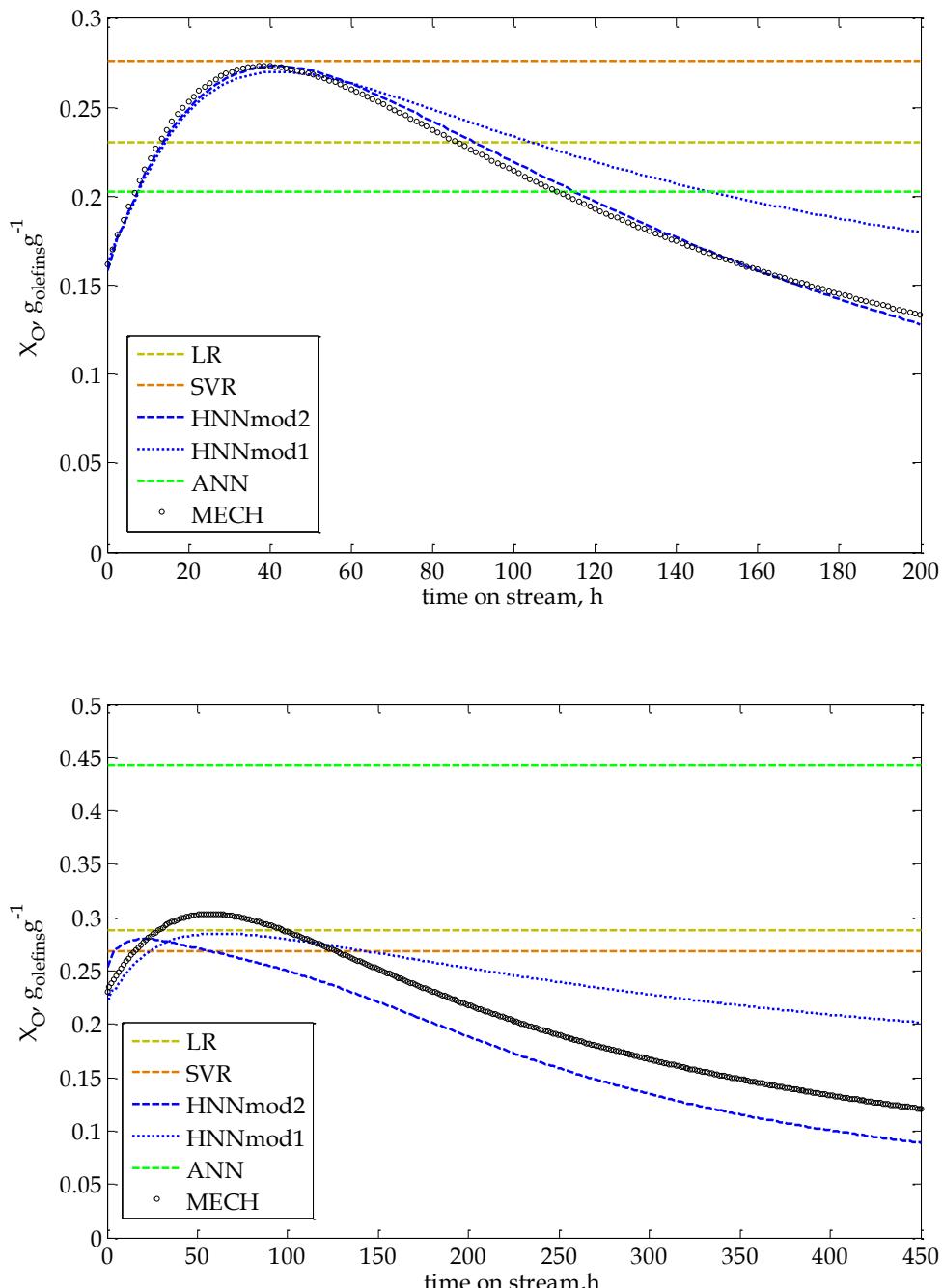


Figure 59. Comparison of the presented model's performance under new operational conditions with different prediction horizons.

5.3. Model Validation

Once defined and selected the most suitable model structure and hyperparameters, a final analysis of the goodness-of-fit of the selected HNN model is performed. Due to the manner it is trained, the errors of the HNN model cannot follow a Gaussian distribution. In fact, the p-value of a Shapiro-Wilk test is 1.028826e-10. Please note that the alternative hypothesis of the Shapiro-Wilk test is that errors do not follow a Gaussian distribution. On these conditions, an F test should not be used to assess the goodness-of-fit of the model. The Kolmogorov-Smirnov test (K-S) (Sheskin 2007) has been selected due to its simplicity and availability. Table 23 presents the results of this test and the mean estimation error of the olefins conversion rate for both the real data and the simulated test dataset.

Table 23. Results of the goodness-of-fit analysis of the HNN model.

| HNN MOD 2 | K-S test | RMSE (golefins $^{-1}$) | MAPE (%) |
|------------------------------|----------|--------------------------|----------|
| Experimental data | 0.1016 | 0.032520 | 15.9145 |
| Simulated data (MECH) | 0.7510 | 0.006634 | 3.6343 |

The null hypothesis of the K-S test (two-sample case) is that both samples in comparison belong to the same distribution. For the MOD 2 model, and for both real experimental data and simulated data, the results of the K-S test do not reject the null hypothesis. Therefore, it cannot be stated that both samples are significantly different. However the results are not conclusive and it is not possible to assure that there are no differences between the distributions.

To assess the goodness-of-fit of the Hybrid model, it should be compared if the model estimates agree both with the result of the Knowledge Model and the Real Measurements. To this end, a Two One-Sided Test (TOST) approach (Robinson et al. 2005) will be used. TOSTs method works by building a region of similarity using expert knowledge, or other procedures

based on scientific knowledge, where both samples could be considered equivalent (Walker & Nowacki 2011). In this particular case, the region of similarity is naturally defined by the experimental error.

After the region of similarity is defined, two one side confidence intervals for a statistic are tested whether they lie within the region of similarity. In this case, the null hypothesis is rejected and it can be concluded that both samples come from the same population. The use of this method provides more trustworthy results than other tests like Kolmogorov-Smirnov or Anderson-Darling. Please note that the null hypothesis of these methods is that two samples are drawn from the same population. Therefore, they can only assess the dissimilarity of the two samples but never the similarity. On the other hand, as it has been explained before, TOST methods have as null hypothesis that the samples come from different populations allowing to conclude the existence of the equivalence.

The TOST method has been used to test the equivalence between the estimations generated by the HNN MOD 2 model and the MECH knowledge model (simulated data). Figure 60 shows the region of similarity and the confidence intervals of the equivalence test for real measurements and for the test simulated dataset. In both cases the confidence intervals are inside the defined region of similarity. Therefore, it can be concluded that the estimations generated by MOD 2 model are equivalent to those generated by the MECH model for both real and simulated data.

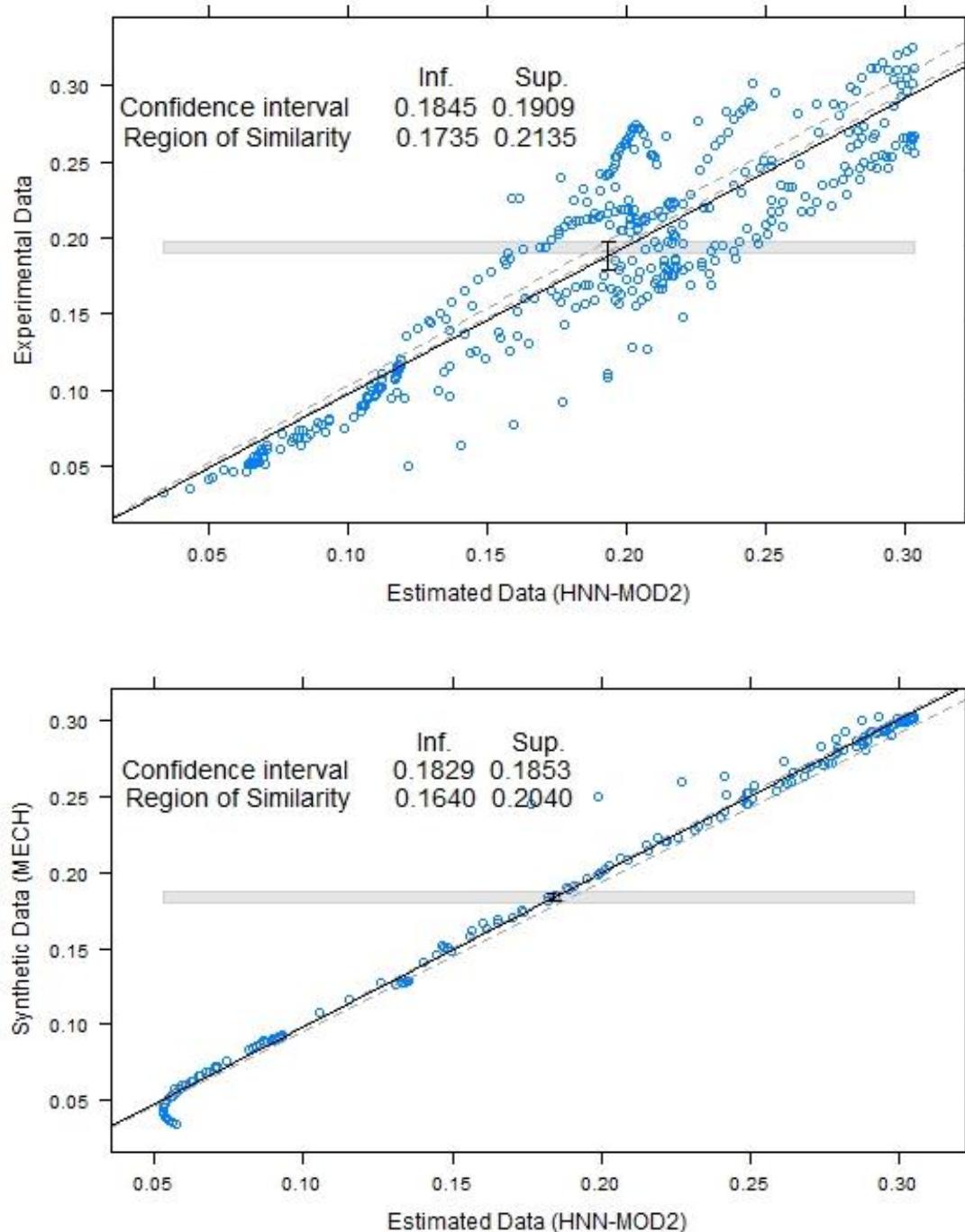


Figure 60. Region of similarity and Confidence interval of the TOST test for Real Measurements (top panel) and for the Test simulated data (bottom panel).

5.4. Conclusions

The modelling of the BTO process has been carried out following the defined methodology and using several modelling techniques. With each of them, several model structures, internal topologies and model hyperparameters have been tested during the modelling procedure.

Figure 61 shows the neural network structure fixed using the modelling methodology as the best model for the BTO process. It uses as inputs the main operating conditions of the process (T , X_w , W/F_{EO}) and the actual and previously calculated catalyst activity levels (a) to compute the main kinetic reaction of the BTO process and calculate the output of the model, which is the olefins conversion rate (X_o). Therefore, there is a hybrid model that includes the knowledge model of the deactivation kinetic to provide the neural model, which will resolve the main kinetic reaction of the process, with information about the catalyst activity state at each time step. The hybrid modelling approach has shown to be the best option due to the influence of the catalyst deactivation, especially for long-term estimations.

The selected model (HNN MOD 2) is equivalent to the mechanistic model (MECH), being its estimations indistinguishable from the real experimental data or from the simulated data generated by the MECH model. HNN MOD 2 (7-20-1) presents a coefficient of determination of 0.9918 for the test dataset and a discrepancy with the mechanistic model, which is the contrast method, of 0.1259 percentage points.

It can be concluded that the HNN MOD 2 (7-20-1) model exhibits very similar performance to the mechanistic model while requiring much lower computational cost. The HNN MOD 2 (7-20-1) model presents estimation times more than 50 times lower than the mechanistic model.

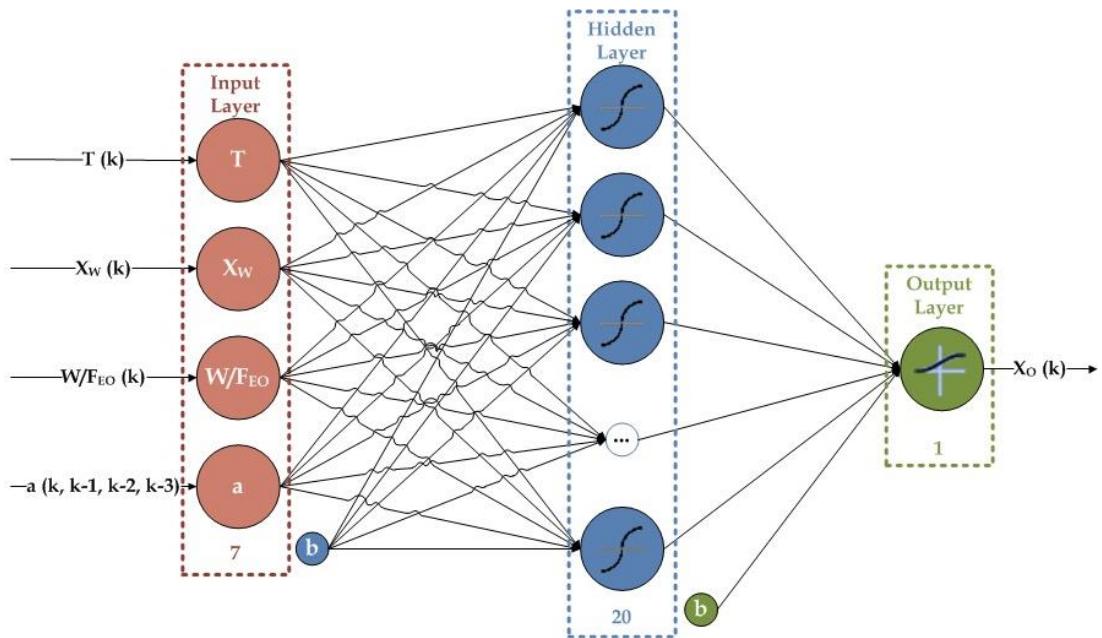


Figure 61. Feed-Forward NN topology of the hybrid model MOD 2 for main kinetic reaction of the BTO process.

OPTIMISATION STRATEGIES

In the following subchapters the results obtained for each optimisation strategy are presented. Additionally, two different models will be used during the optimisation procedure in order to simulate the behaviour of the BTO process. The results obtained using the mechanistic and a neural model (used as surrogate model) of the process will be compared to analyse if the last one is able to reach to similar results than as those obtained with the computationally more expensive mechanistic model. The surrogate neural model used in the optimisation is the previously developed Feed-Forward ANN based hybrid model that combines the kinetic model of the catalyst deactivation, used to calculate the evolution of the catalyst activity, with an ANN that describes the process kinetics to obtain light olefins.

6.1. Optimisation results

6.1.1. Optimisation at zero time of stream

As previously exposed, the optimisation at zero time of stream has been carried out using the equation (31) as the cost function of the optimisation algorithm.

Figure 62 shows the behaviour of the BTO process under the best operational conditions found in order to maximise the conversion rate of olefins at zero time on stream. Note that this optimisation procedure, does not take into account the catalyst deactivation in the search for the best operational conditions. As can been seen, due to a faster catalyst deactivation, the production stopping criteria (t_f) is reached in a relatively early time when comparing with the solutions obtained with fixed-shape temperature trajectories or with dynamic trajectories for each operating

variable. The temperature is near to the superior limit in order to avoid an irreversible deactivation of the catalyst. This provoke a maximum conversion rate but with a fast catalyst deactivation. In contrast, the water content in the feed (X_w) is minimal which also explain the rapid catalyst deactivation, while the space-time (WF_{E0}^{-1}) presents also a reduced value according to the selected temperature and water content.

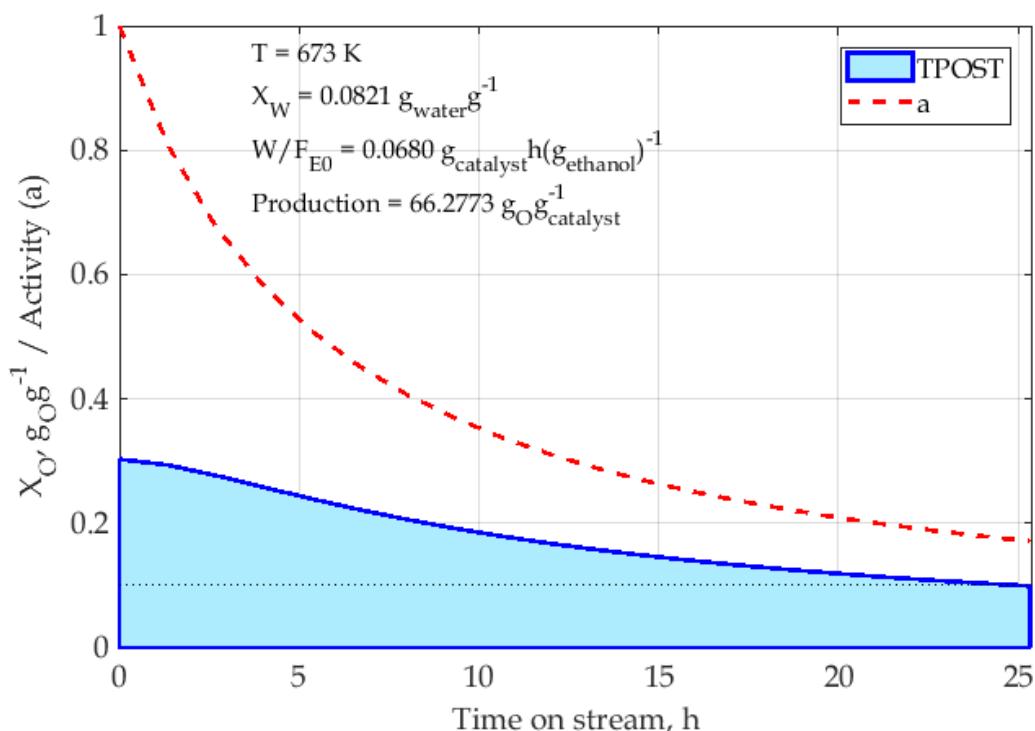


Figure 62. BTO process performance under the operational conditions obtained when optimising the conversion rate of the olefins at zero time on stream.

The maximum production reached with these operation conditions is $66.27 \text{ g}_O \text{ g}_{\text{catalyst}}^{-1}$ at the cost of deactivating the catalyst very quickly. This effect will increase the catalyst regeneration costs and the risk of the irreversible deactivation of it.

6.1.2. Constant and dynamic optimisation

The first and second optimisation strategies consist on finding the optimum constant set-points and dynamic trajectories of the temperature profiles.

Table 24 and Table 25 summarizes the average and best TPOST (results of Equation (20)) for the main temperature trajectories (Equations (21)–(29)) and for constant set-points (f_0) obtained using the mechanistic and the neural surrogate models respectively. Being the optimisation procedure stochastic, the optimisation algorithm with each trajectory type has been launched 50 times to guarantee its convergence.

At first sight, near all proposed trajectory functions achieve noticeable better results than the contrast method (trajectory f_0). In both cases the best trajectory seems to be f_8 and f_9 closely followed by trajectories f_1 . The standard deviation of the production is quite low in near all cases, when comparing to the standard deviation of the standard normal distribution, suggesting that the method converges systematically to a single set of operational conditions. When comparing both tables, these results are repeated independently to the model used to simulate the BTO process during the optimisation procedure. Although, there are slight differences in the absolute results for each trajectory profile type, the main conclusions comparing the trajectories are equal.

Table 24. Comparison of the mean production results for the best operating conditions when repeating several times the evolutionary optimisation using the MECH model.

| MECH | TPOST ($\text{g}_{\text{og}}^{-1}\text{catalyst}$) | | |
|----------|--|---------------|----------|
| | Best TPOST | μ TPOST | σ |
| $f_0(t)$ | 73.761 | 70.259 | 2.381 |
| $f_1(t)$ | 86.584 | 81.911 | 0.830 |
| $f_2(t)$ | 82.509 | 80.596 | 1.271 |
| $f_3(t)$ | 81.548 | 78.030 | 2.227 |
| $f_4(t)$ | 75.591 | 73.346 | 1.891 |
| $f_5(t)$ | 82.004 | 80.347 | 1.347 |
| $f_6(t)$ | 78.861 | 74.206 | 1.758 |
| $f_7(t)$ | 82.380 | 81.053 | 0.782 |
| $f_8(t)$ | 89.973 | 88.215 | 1.412 |
| $f_9(t)$ | 88.670 | 86.287 | 1.731 |

Table 25. Comparison of the mean production results for the best operating conditions when repeating several times the evolutionary optimisation using the HNN surrogate model.

| HNN | TPOST ($\text{g}_{\text{og}}^{-1}\text{catalyst}$) | | |
|----------|--|---------------|----------|
| | Best TPOST | μ TPOST | σ |
| $f_0(t)$ | 69.963 | 66.101 | 2.832 |
| $f_1(t)$ | 83.247 | 80.477 | 0.891 |
| $f_2(t)$ | 81.086 | 79.806 | 0.853 |
| $f_3(t)$ | 79.994 | 75.523 | 2.286 |
| $f_4(t)$ | 72.493 | 69.516 | 1.607 |
| $f_5(t)$ | 82.139 | 80.108 | 1.540 |
| $f_6(t)$ | 65.837 | 65.596 | 0.106 |
| $f_7(t)$ | 72.157 | 69.476 | 1.689 |
| $f_8(t)$ | 86.578 | 84.742 | 1.477 |
| $f_9(t)$ | 86.747 | 82.424 | 6.313 |

Figure 63 and Figure 64 use a Notched Box Plots (John M. Chambers 1983) to present the best solution of each global iteration for each trajectory type (constant, polynomial, exponential, logarithmic and trajectories inverse to the activity). The boxes represent for those trajectories the first and third quartile of the production of olefins (TPOST). The central line crossing the boxes represents the median value and the whiskers represent 1.5 times the interquartile distances. The white points represent outlier values.

The notches around the median value represent a 95% confidence interval. So, if two notches do not overlap, there is “strong evidence” that those median values are statistically different (John M. Chambers 1983). Following this rule, it seems that trajectories inverse to catalyst deactivation (f_8 and f_9) provide the best options to maximize the total production. Although, given that their notches overlap, they might be indistinguishable one from the other. Anew, the conclusions obtained following this rule, is the same independently to the model used during the optimisation. In both cases, f_8 and f_9 provided the best solutions, closely followed by the f_1 trajectory. This trajectory profiles (f_1) provide solutions slightly worse but with a significantly lower variability.

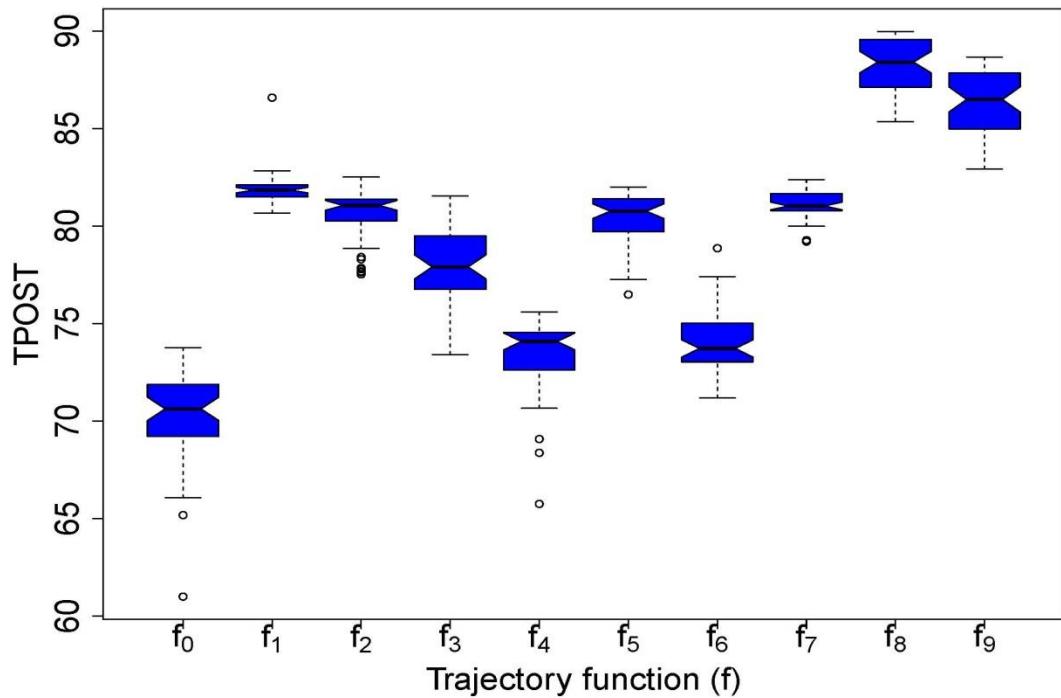


Figure 63. Comparison between the different results for the proposed temperature trajectory functions using the MECH model.

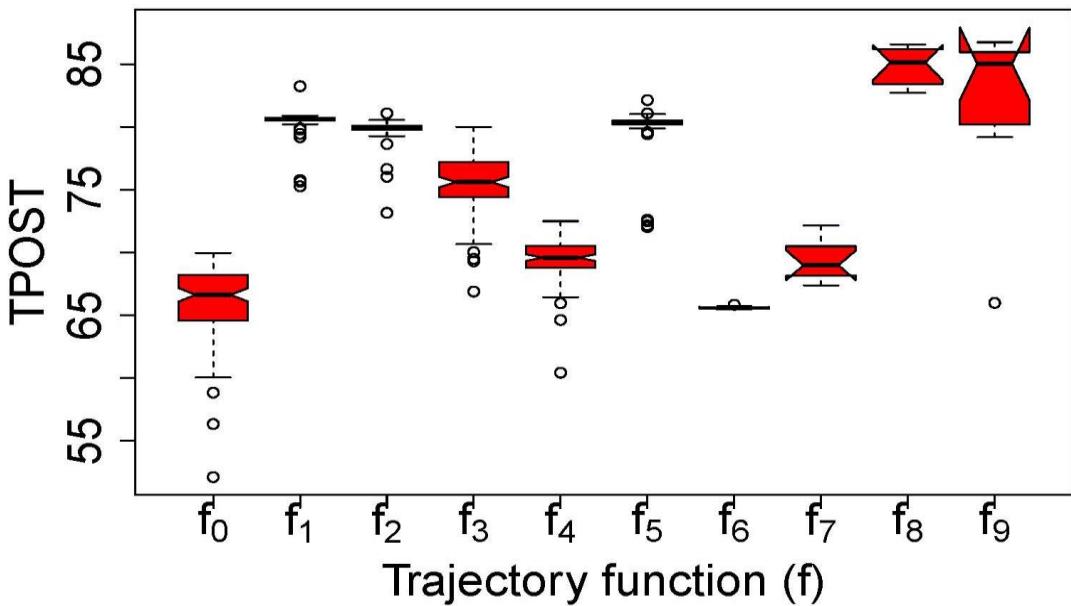


Figure 64. Comparison between the different results for the proposed temperature trajectory functions using the HNN surrogate model.

In order to fully assess the operational trajectories that achieve the highest olefins production, it is needed to make a Multiple Comparison Test as ANOVA with Tukey's HSD post-hoc. Please note that pairwise comparison should not be made in order to avoid the Multiple Comparisons Problem (Demšar 2006; Garcia & Herrera 2008). To do a Multiple Comparison Test, first a Shapiro-Wilk test (S-W) should be made in order to assess if a parametric test (as an ANOVA test) could be made.

Table 26 shows that some of the productions of olefins variables do not follow a Gaussian distribution (those with score lower than 0.05) and therefore a non-parametric test should be used. Given that the experimental values obtained from several repetitions are paired, a Friedman test with Nemenyi post-hoc should be used (Demšar 2006; Myles Hollander et al. 2014).

Table 26. Results of the Shapiro-Wilk test over the production of olefins for each trajectory.

| Trajectory | S-W (MECH) | S-W (HNN) |
|------------|-------------------------|------------------------|
| $f_0(t)$ | 0.0012 | 7.297×10^{-8} |
| $f_1(t)$ | 4.8550×10^{-9} | 2.2×10^{-16} |
| $f_2(t)$ | 1.4420×10^{-6} | 2.2×10^{-16} |
| $f_3(t)$ | 0.0876 | 0.0012 |
| $f_4(t)$ | 7.8487×10^{-7} | 7.456×10^{-8} |
| $f_5(t)$ | 0.0005 | 2.2×10^{-16} |
| $f_6(t)$ | 0.0007 | 0.2902 |
| $f_7(t)$ | 0.0929 | 0.3300 |
| $f_8(t)$ | 0.0033 | 0.1970 |
| $f_9(t)$ | 0.0098 | 5.17×10^{-4} |

Figure 65 graphically represents the results of the Friedman test with Nemenyi post-hoc. As can be seen, the p-value of the Friedman test is significant (when comparing all the trajectories) so that there is at least one trajectory that is statistically different from the others. This fact is repeated

for both optimisation procedures, using the MECH model (left panel) and with the HNN neural model (right panel).

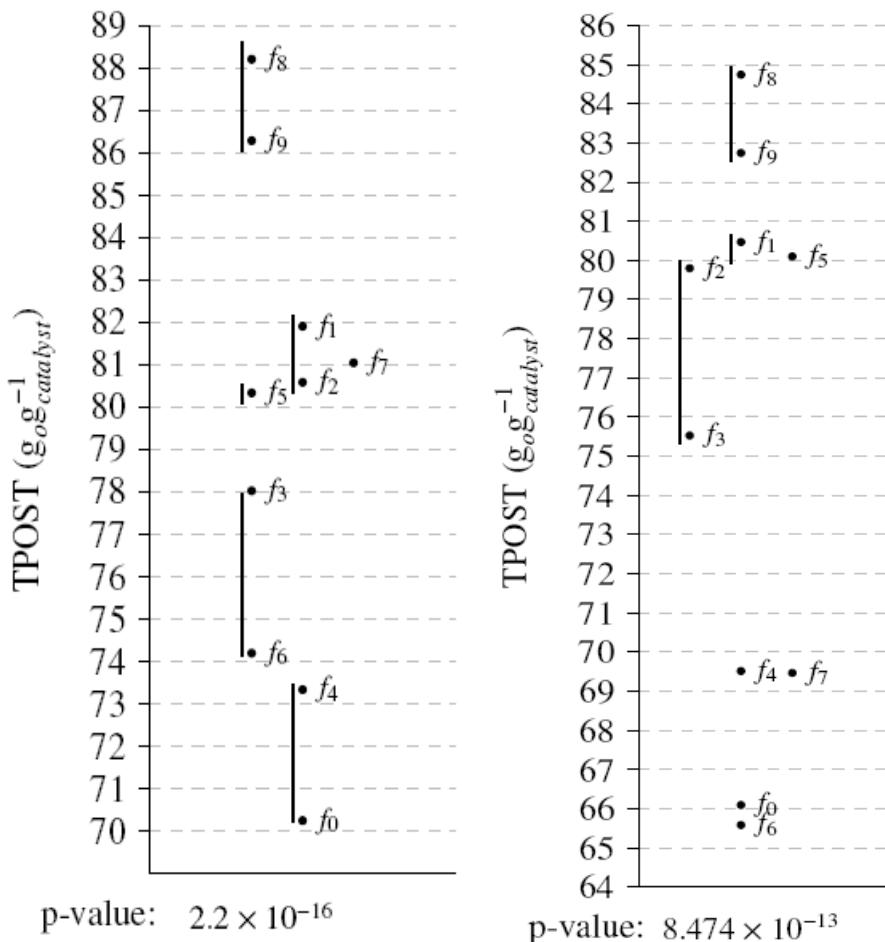


Figure 65. Results of the Friedman test (p-value) and the Nemenyi post-hoc over the different trajectories (Left panel: MECH model; Right panel: HNN model). Linked are the trajectories that are not statistically different attending to the results of the post-hoc test.

The Nemenyi post-hoc presents quite similar results in both cases. On the left panel five main groups can be observed. f_0 and f_4 are grouped together, being statistically indistinguishable. Next, solutions provided by f_6 and f_3 form a second group, the f_5 form a group by itself and f_1 , f_2 , f_7 are grouped together. And finally, the best operational conditions are provided by the f_8 and f_9 trajectories, with results that are statistically indistinguishable. On the right panel, three main groups can be observed. Anew, f_8 and f_9 form a

group by themselves, being significantly different from the other trajectories and obtaining the best results. Next, solutions f_1 and f_5 , and f_2 and f_3 form two differentiable groups respectively. Finally, the poorer results are obtained by the trajectories f_4 , f_7 , f_0 and f_6 , which form all them a group by itself. These results confirm the previously exposed observation. Trajectories inverse to the catalyst deactivation rate (f_8 and f_9 trajectories) obtain the best results and there are no evidences of the existence of significant differences between them. Additionally, linear trajectories (f_1) present closed results to the previous ones, but statistically significantly worse than them.

Finally, Figure 66 and Figure 67 show the behaviour of the process when using the trajectory inverse to the catalyst deactivation f_8 (Equation (28)). This is the best solution obtained by the optimisation procedure using fixed-shape temperature profiles. The obtained best solutions with both models (MECH in Figure 66 and HNN in Figure 67) are quite similar; with the same temperature profiles and similar values for all operating variables of the process. The maximum production (TPOST) has been $89.97 \text{ g g}^{-1}\text{catalyst}$ for the MECH model and $86.57 \text{ g g}^{-1}\text{catalyst}$ for the HNN model. The temperature profiles slowly grow while the activity slowly decreases reaching the upper bound quite late in time (more than 700 h). Moreover, the amount of feed water (X_w) is quite high. These are the main factors that allows the production as long as 797 h (729 h using the neural model in the optimisation) due to the fact that water competes with the coke for the active centres of the catalyst attenuating the deactivation. Finally, the space-time ($WF^{-1}\text{EO}$) is moderately high looking for the maximum conversion of olefins for that particular conditions of water and temperature variables. This particular happens due to the fact that the conversion reaches a maximum in that particular space-time as it is an intermediate product in the reaction mechanism. In the same way, the best production obtained using f_9 trajectory (Equation (29)) is obtained with similar operational conditions. Using convex temperature profiles which slowly grows, and using high amount of water in the feed in order to extend the catalyst lifespan.

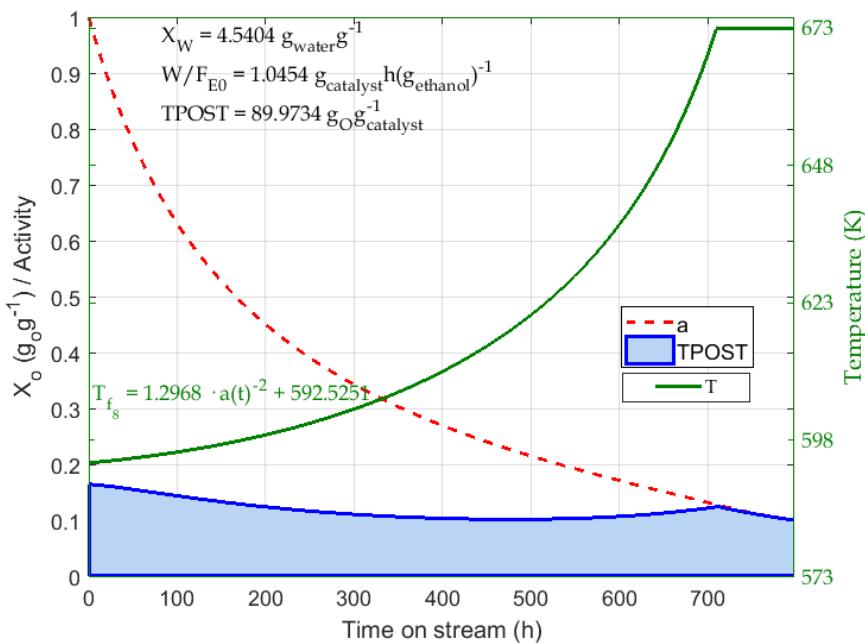


Figure 66. BTO process performance under the best solution (f_8) provided by the evolutionary optimisation using temperature trajectories (MECH model).

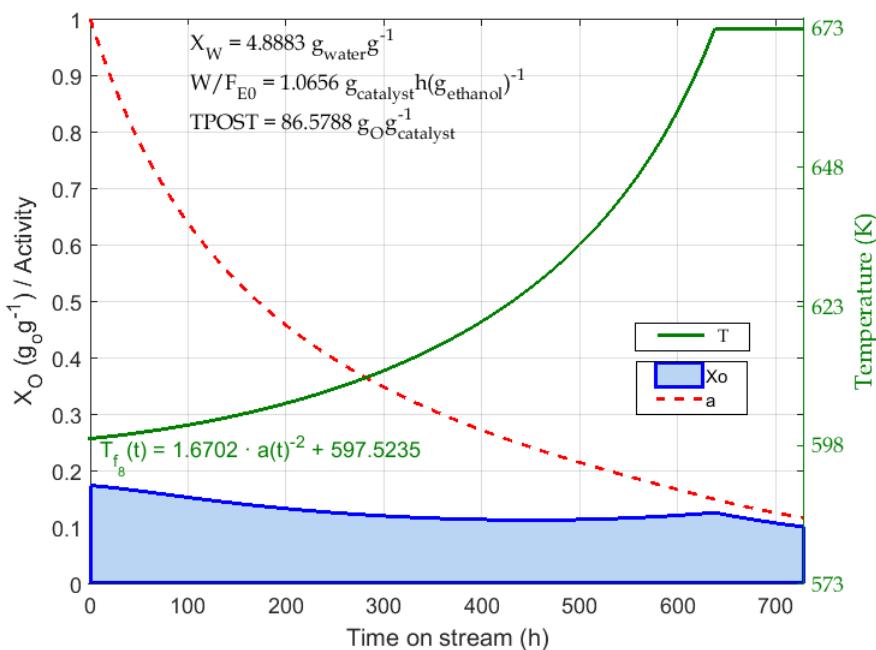


Figure 67. BTO process performance under the best solution (f_8) provided by the evolutionary optimisation using temperature trajectories (HNN model).

6.1.2.1. Conclusions

The results for the dynamic optimisation of the temperature trajectories seem to be conclusive. The best solution capable to maximise the production of olefins, while extending the catalyst lifespan, are the temperature trajectories inverse to the catalyst deactivation. The slow increase of the temperature counteracts the loss of the catalyst activity and therefore the desired production of olefins. At the same time, in order to avoid an early fast catalyst deactivation, the water content in the feed is near to the maximum.

Additionally, it has been demonstrated that the use of surrogated neural models during the optimisation procedures is very satisfactory. The neural structure is able to optimise the process estimating the dynamic behaviour of the process even for long-term prediction scenario. The productions estimated by the neural models are slightly lower than those obtained by the mechanistic model. However, the obtained conclusion and the best operational conditions are equivalents to those obtained using the contrast method (MECH model based optimisation, but with a significantly lower computational cost (up to 50 times lower).

Therefore the presented optimisation methodology is valid to search for the optimum areas with a considerably lower computational cost.

6.1.3. Dynamic neural optimisation

A third optimisation strategy has been tested. This consists on optimising dynamically all the operating variables (T ; X_w ; W/F_{E0}) and not only the temperature profiles. Due to the computational complexity of the previously used optimisation procedure, in this case, the previously presented ANN based approach has been selected.

Table 27 shows the results obtained with different number of hidden neurons to generate the trial profiles for the three operating variables. Anew, being the optimisation procedure stochastic, the whole optimisation procedure has been launched 50 times to guarantee its convergence.

The previously presented solutions have been obtained with fixed-shape temperature profiles, and those really improve the final production. In this case, using the operational profiles directly generated by the evolved neural network, the results are even better. The results with 5 or more hidden neurons clearly overcome, at first glance, the previous obtained results. Nevertheless, the obtained standard deviation seems to be quite high, so deeper analyses are needed.

Table 27. ANN training results for the implemented neural structures.

| Hidden neurons | TPOST ($\text{g}\text{og}^{-1}\text{catalyst}$) | | |
|----------------|---|---------------|--------------|
| | Best TPOST | μ TPOST | σ |
| n = 4 | 101.425 | 86.693 | 8.436 |
| n = 5 | 98.329 | 87.547 | 7.350 |
| n = 6 | 101.729 | 94.825 | 4.591 |
| n = 7 | 100.52 | 90.615 | 7.686 |
| n = 8 | 101.585 | 89.403 | 6.729 |

Figure 68 shows a Notched Box Plot to present the obtained solutions of each global iteration during the optimisation procedure. It seems that the trajectories provided by the ANN structures with 5 hidden neurons or more are statistically similar between them and with relatively high standard deviations which increases the size of the represented notches approaching to worse solutions.

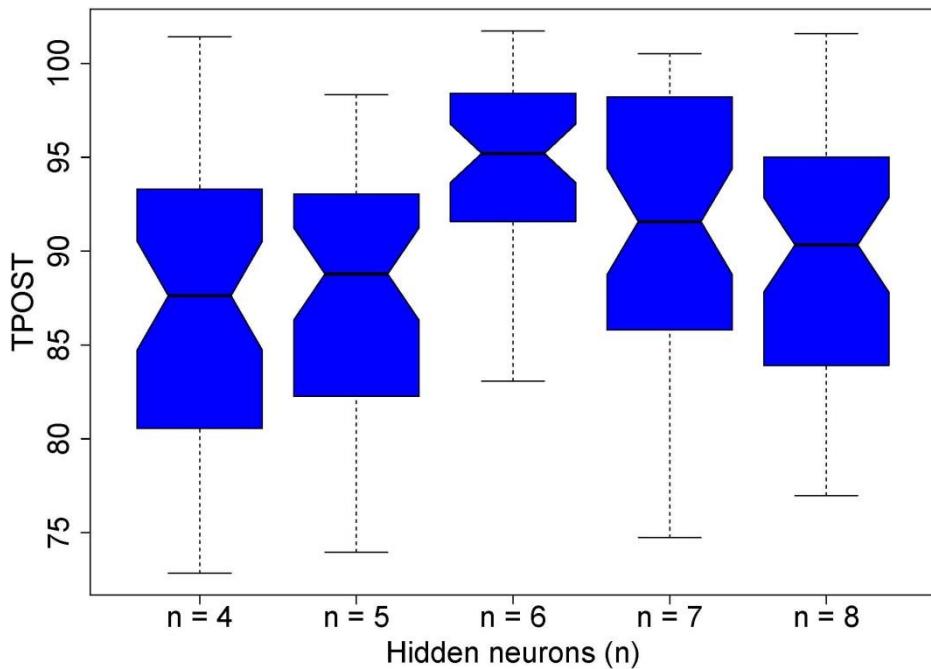


Figure 68. Comparison between the different results for the proposed neural structures.

Anew, the same previously used procedure will be used to fully assess the optimisation strategy that obtained better solutions. First, a Shapiro-Wilk test should be made in order to assess if a parametric test (as an ANOVA test) could be made. Table 28 shows that at least one of the solutions does not follow a Gaussian distribution. Therefore the Friedman test with Nemenyi post-hoc will be used. Moreover, the solutions obtained with f_8 and f_9 temperature profiles will be included in this test in order to analyse if the new solutions obtained using ANN are statistically distinguishable from the previously obtained solutions for fixed-shape temperature profiles.

Table 28. Results of the Shapiro-Wilk test over the production of olefins for the proposed neural structures.

| Trajectory | |
|-------------|--------|
| n= 4 | 0.0091 |
| n= 5 | 0.0100 |
| n= 6 | 0.0959 |
| n= 7 | 0.0012 |
| n= 8 | 0.1624 |

Figure 69 graphically represents the results of the Friedman test with Nemenyi post-hoc. As can be seen, the p-value of the Friedman test is significant (when comparing the trajectories f_8 and f_9 , and those profiles generated by the ANN) so there is at least one solution that is statistically different from the others. The Nemenyi post-hoc presents three differentiable groups. The ANN with 4 hidden neurons (n_4) and the trajectory f_9 form a first group. The ANN structures with 5, 7 and 8 hidden neurons (n_5 ; n_7 ; n_8) which results are all of them statistically indistinguishable from each other and also indistinguishable from f_8 , form the second group. This second group overcome the results of the previous one (n_4 ; f_9). And the best operational conditions are provided by the ANN structure with 6 hidden neurons which results are statistically different from the others.

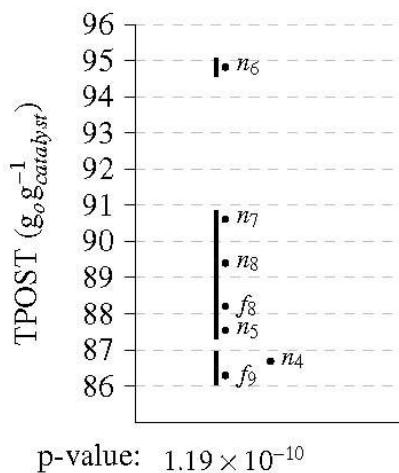


Figure 69. Results of the Friedman test (p-value) and Nemenyi post-hoc over the solutions generated by ANN and trajectories f_8 and f_9 . Linked trajectories are not statistically different attending to the results of the post-hoc test.

In order to fully assess that when only optimising the temperature trajectories, that are solutions as f_8 , are able to obtain solutions indistinguishable or not from those obtained by ANN with 6 hidden neurons, a deeper analysis has been carried out. The complete optimisation procedure for both of them has been repeated 150 times, and a new statistical analysis has been performed. Figure 70 shows graphically in a Notched Box Plot that both solutions seem to be statistically different. In this case, to compare the two samples, the Wilcoxon signed rank test has been selected (David F. Bauer 1972; Myles Hollander et al. 2014). The result (p-value $< 2.2000 \times 10^{-16}$) confirms that the dynamic solutions obtained by the ANN with 6 hidden neurons are significantly better than those obtained optimising only the temperature profiles.

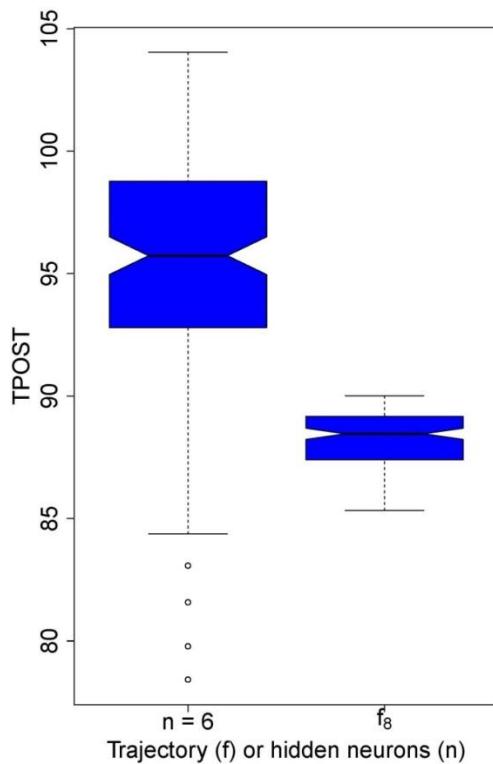


Figure 70. Comparison between the different results obtained with ANN (n_6) and f_8 temperature profiles.

Finally, Figure 71 shows the behaviour of the BTO process (Figure 71a) under the best solution found by the evolutionary optimisation using an ANN to generate the operational trajectories (Figure 71b, Figure 71c and Figure 71d). The best solution was found using the third optimisation strategy with 6 neurons in the hidden layer of the ANN. The maximum TPOST has been $101.729 \text{ g}_{\text{O}_2}^{-1} \text{ catalyst}$. This solution provides a 37.91 % higher olefins production than using constant set-points, which is one of the reference operational strategy. A slow increase of the temperature profile at the begining and the relatively low ranges of temperature over the whole operation are able to maintain a low deactivation rate. Moreover, a high amount of water in the feed (X_w), counteracts the catalyst deactivation. These two aspects are the main factors that maximize the catalyst lifespan and extend its life for more than 500 h. Moreover, these slow changes in the temperature have two good operational properties: the amount of power needed is bounded and avoids the occurrence of hot spots in the catalyst.

Moreover, the space-time (WF^{-1}_{E0}) grows up to obtain the maximum conversion rate of olefins in the reactor output for that particular condition of water and temperature variables. In the same way, the best solutions obtained using f_8 or f_9 trajectories, although being slightly worse (productions (TPOST) of 89.97 and 88.67 $g_0 g^{-1}_{catalyst}$ respectively), are able to counteract the effect of the catalyst deactivation extending the production steps.

The evolved neural networks have been able to find different trajectory types for each operational variable that maximize the production objective. These trajectories have been able to detect simultaneous regimen changes in all the operational variables (as that observed in Figure 72Figure 71 at time 103 h) that allow an increase of the production. These necessary regimen changes in the trajectories of the control variables are very difficult to detect using standard techniques with no previous knowledge about the process. On the other hand, ANN training techniques show a relatively high standard deviations (Table 27) compared with the second optimisation strategy (Table 24). This can be explained due to the fact that the search space is infinitely higher than in the case of using fixed-shape temperature trajectories.

These regimen changes, that allow the maximisation of the production, are also seen in the best solutions provided by ANN based optimisation. Figure 72 shows the behaviour of the process under the best solution provided by the ANN with 8 hidden neurons. The production profile is similar to the previous one with a regimen change after 220 hours of production. In this case the effect is clearly notice in all the operating variables, which change their previously followed trend. Once again, the water content in the feed start at its maximum possible value in order to counteract the catalyst deactivation. And the space-time grows up according to the other process variables. Finally, the temperature profile is completely different to the previous one (Figure 71) but maintaining the average within a similar range.

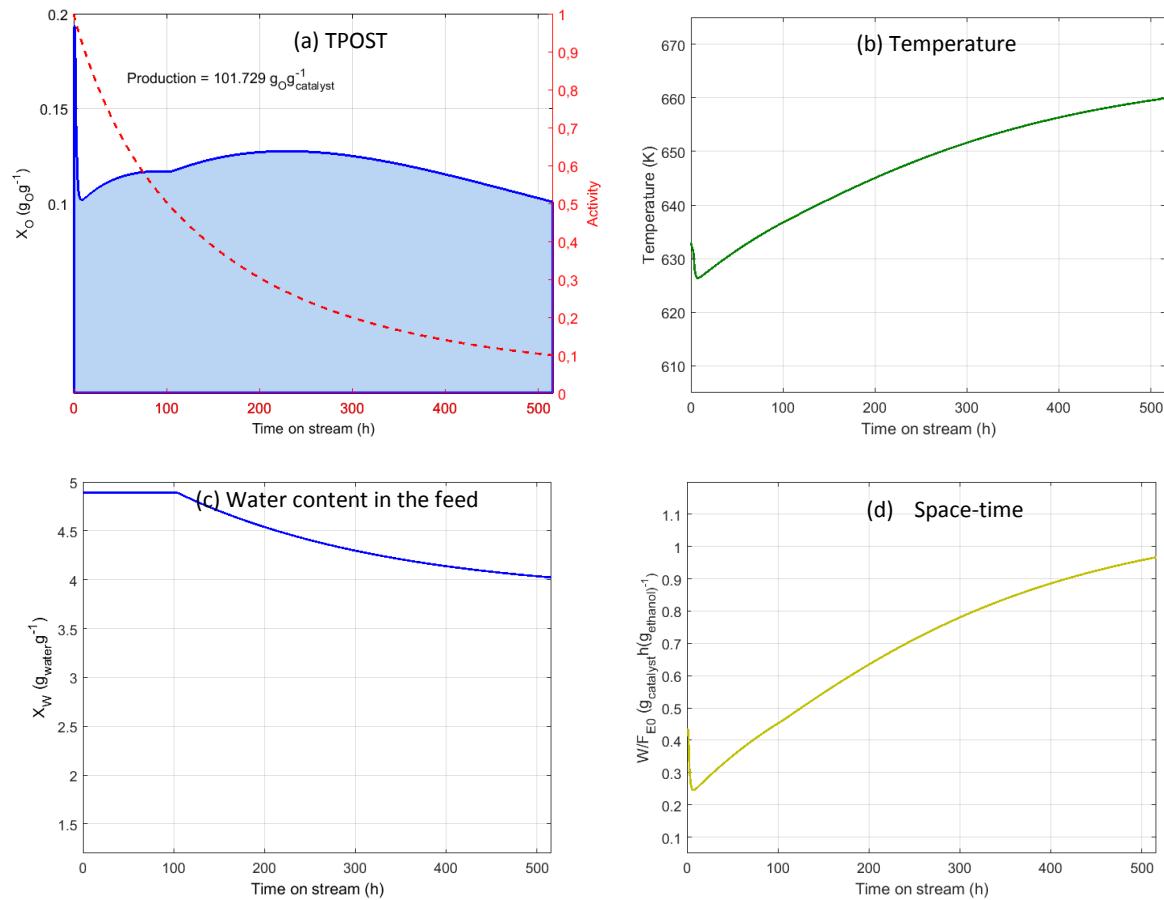


Figure 71. BTO process performance and trajectories of the operational conditions for the best solution provide by the evolutionary optimisation ($n = 6$).

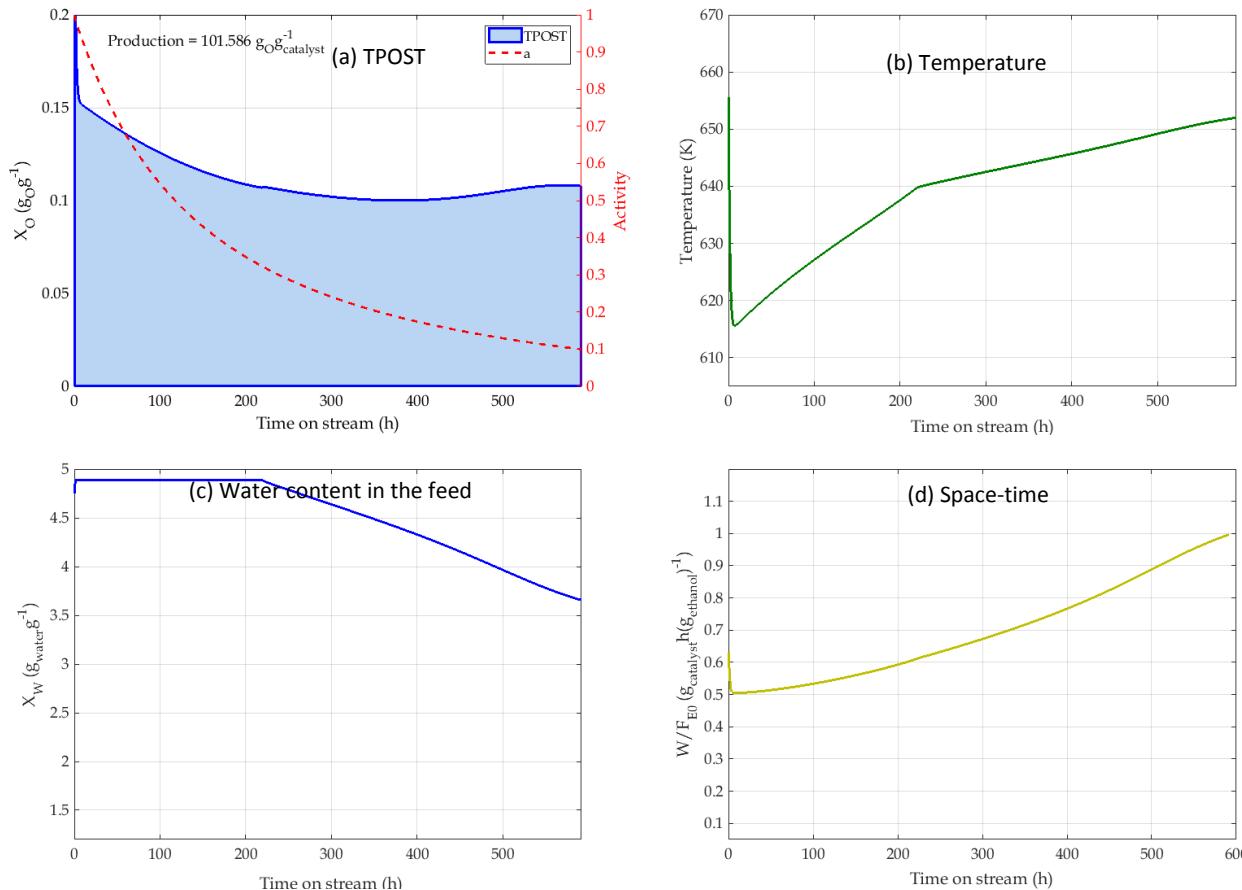


Figure 72. BTO process performance and trajectories of the operational conditions for and alternative solution provide by the evolutionary optimisation ($n = 8$).

6.2. Comparison and discussion

As previously stated, the standard optimisation procedure to study the best operational conditions for catalytic processes will be used as contrast method for the proposed optimisation strategies (see section 4.3.4).

The best operation at zero time of stream (see section 6.1.1) provoke a very fast catalyst deactivation which may increase the total production costs, as well as increasing the effect of the irreversible deactivation of the catalyst.

Figure 73 compares the best solution obtained with each optimisation strategy. The optimisation carried out at zero time on stream (blue dotted line) is able to obtain $66.27 \text{ g} \cdot \text{g}^{-1} \text{ catalyst}$, reaching the stopping criteria in 25.3 hours. In contrast the constant optimisation (orange line), which has been carried out taking into account the catalyst deactivation, reaches to the stopping criteria in 49 hours, but is able to improve the previous production in a 11.15 %. The solutions provided by both dynamic (black line) and fixed-shape temperature trajectories (green dotted line), are able to extend the catalyst lifespan up to 500 hours and 800 hours respectively. However, the maximum production is obtained using the dynamic trajectories of all operating variables generated by the ANN. These operation trajectories are able to obtain a 53.48 % and a 35.75 % higher production than the production obtained with zero time optimisation procedures respectively.

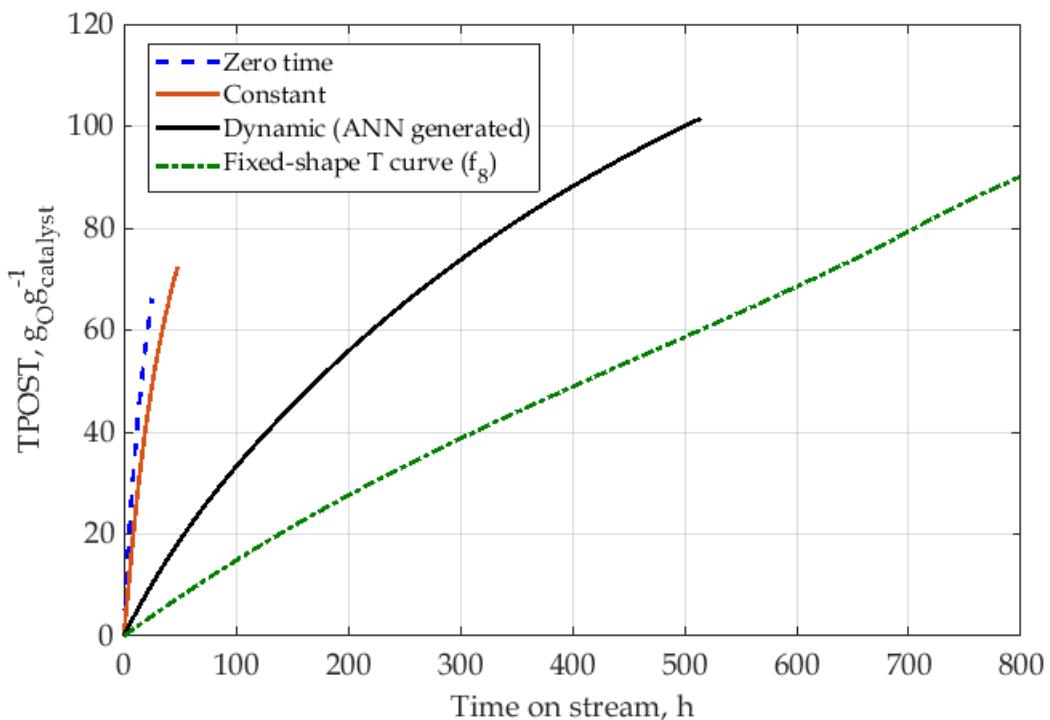


Figure 73. Comparison of the production results obtained with each optimisation strategy.

Therefore, the best solution in order to maximise the TPOST and to extend the catalyst lifespan has been obtained by the trajectories generated by the ANN. However, a key point to optimise the production of olefins will be cost of the needed catalyst and its regeneration and maintenance costs after each production step. Depending on these two costs, it will be more or less interesting to extend the catalyst lifespan in order to continue the production step or to carry out shorter production cycles with more catalyst regeneration processes. Moreover, maintaining constantly the maximum temperature as is needed to maximize the production at zero time of stream, involves higher plant operating costs.

6.3. Study of the operational conditions

The implemented optimisation strategies make special emphasis to the temperature because it is a dominant operating variable. Therefore, during the optimisation procedure the optimisation of the temperature trajectories has been one of the main objectives to maximize the total production of olefins. However, the behaviour of the water content in the feed and the space time of the process have also an important influence in the extension of the catalyst lifespan which is the secondary optimisation objective.

Therefore, to analyse these two operating variables, a grid search study has been performed. The optimisation procedure has been repeated several times using the temperature trajectory f_8 for each fixed value of the water content in the feed using a grid mesh. The grid for the X_w variable has been defined as in the previously used operational limits, between [0.05, 5] and with a step of 0.05 $\text{g}_{\text{water}}\text{g}^{-1}$. For each X_w value of the grid, the optimal parameters of the temperature profile and the optimal WF^{-1}_{E0} values in each case have been searched using the proposed optimisation methodology.

Finally, this study will be used to analyse different optimal solution depending on the catalyst regeneration costs after each production step.

6.3.1. Catalyst lifespan extension

On the one hand the effect of the temperature in the catalyst deactivation rate is clear, even arriving to the irreversible catalyst deactivation when using to high temperatures. Thus, as can be seen during the optimisation procedure, the best operational conditions maintain intermediate values of the temperature, starting with relatively low temperatures (above the minimum necessary for the process to take place the transformation of the bioethanol into olefins), and once the catalyst deactivation has started, the

temperature increases gradually in order to maximise the production, warranting a minimum production of olefins during the production cycle.

On the other hand, a key to extend the catalyst lifespan is the water content in the feed. The presence of water in the feed of the reactor implies that the water molecules compete with coke for the active centres, avoiding the accumulation of coke in the active centres of the catalyst and thus it will attenuate the catalyst deactivation. The optimisation procedure has been carried out maintaining similar temperature profiles while changing the X_W within the grid search. The results show how the average production times for each cycle of the best solutions increases with the X_W . Figure 74 shows how the increment of the water content extends the catalyst lifespan, resulting larger production cycles.

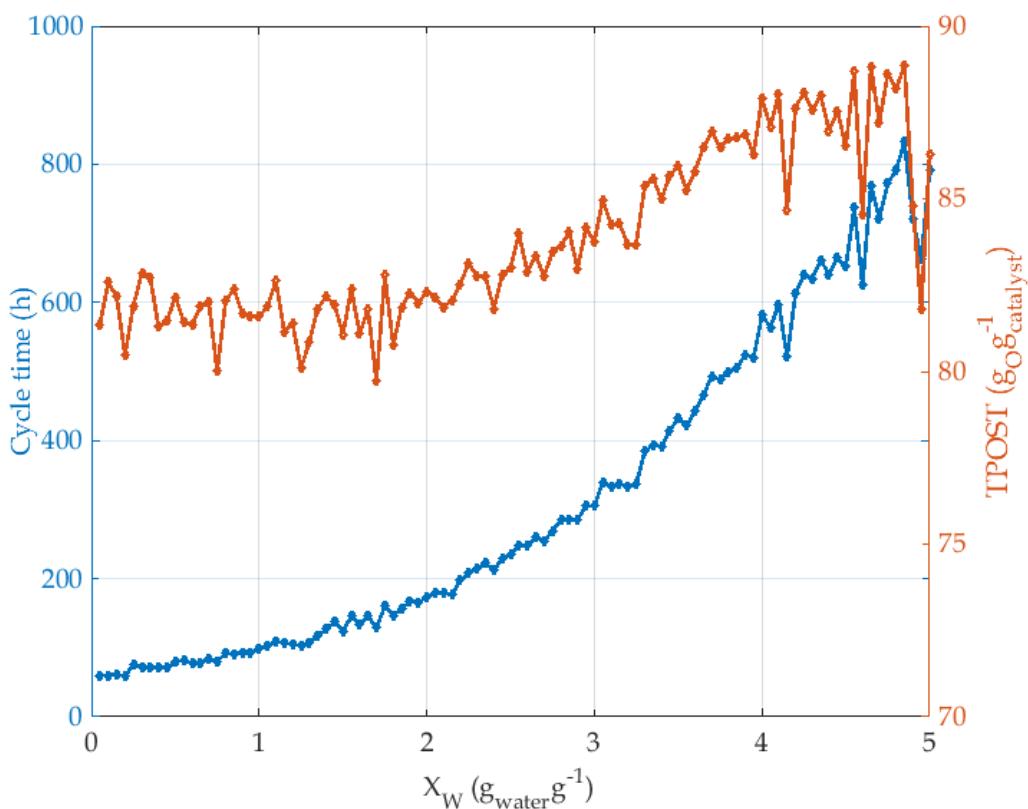


Figure 74. Effect of the water content in the feed in the catalyst lifespan extension.

Additionally, the optimal water content in the feed and space-time show a correlation between them in order to reach the optimisation objectives. Figure 75 shows the distribution of the solutions for each X_w value and its corresponding WF^{-1}_{E0} in order to maximize the production of olefins and extend the catalyst lifespan. As can be seen, there exist a relation between both variables, obtaining the best solutions with a specific combination of these two variables.

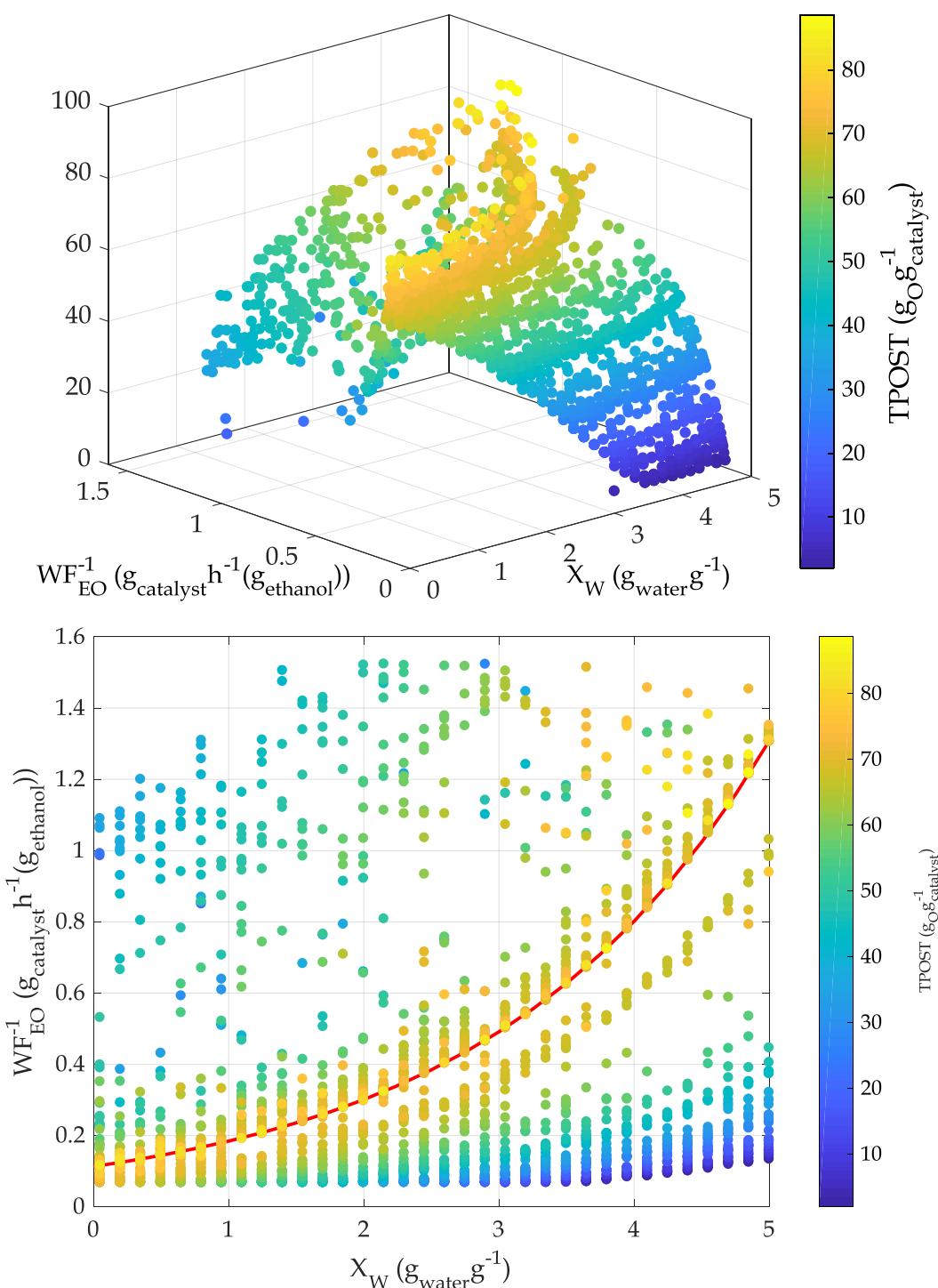


Figure 75. Optimisation results for each fixed X_W value. The red line represents the relation between X_W and the space-time with the maximum TPOST.

The best solutions (with the maximum TPOST) for each X_W values are obtained with space-time values that follow an exponential relation depending on the X_W value. Figure 76 shows that the best solutions (with TPOST higher than $80\text{g}_{\text{catalyst}}^{-1}$) are obtained following the next exponential relation:

$$WF_{E0}^{-1} = a \cdot \exp(b \cdot X_W), \quad (31)$$

being the constant parameters a and b , 0.113 (0.1127, 0.1132) and 0.4898 (0.4893, 0.4903) respectively.

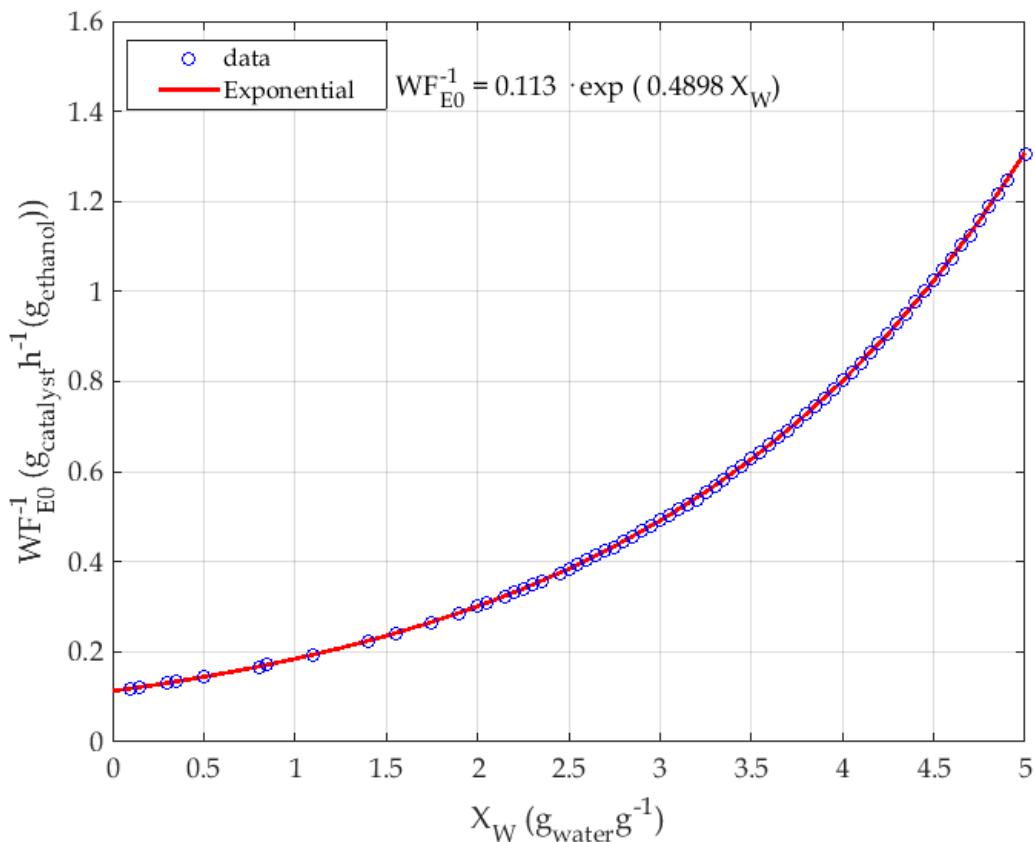


Figure 76. Correlation of WF_{E0}^{-1} values for the best solutions for each X_W values.

The Figure 77 clearly shows that the maximum production value that can be reached for each quantity of water in the feed moves through the possible

values of space-time. These maximum values fit with the previously shown Equation (31) that relates the X_w and the space-time.

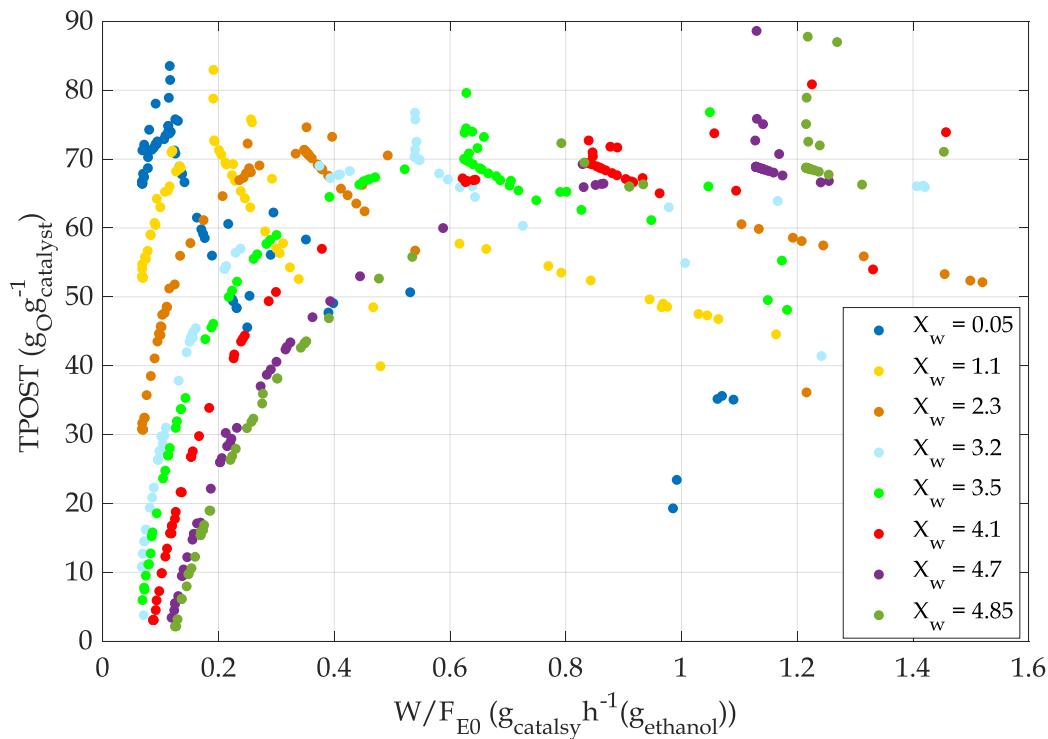


Figure 77. TPOST maximum values for each combination of X_w and $\text{W}/\text{F}_\text{E0}$ values.

Finally, Figure 78 shows the results of performing another optimisation procedure for the temperature profiles, following the exponential relation for each X_w value. For each value of X_w and its corresponding WF^{-1}_E0 value of equation (31), the temperature trajectory is optimised in order to maximise the TPOST. As can be seen, as previously exposed the best solutions are obtained with high values of water in the feed due to the extension of the catalyst lifespan.

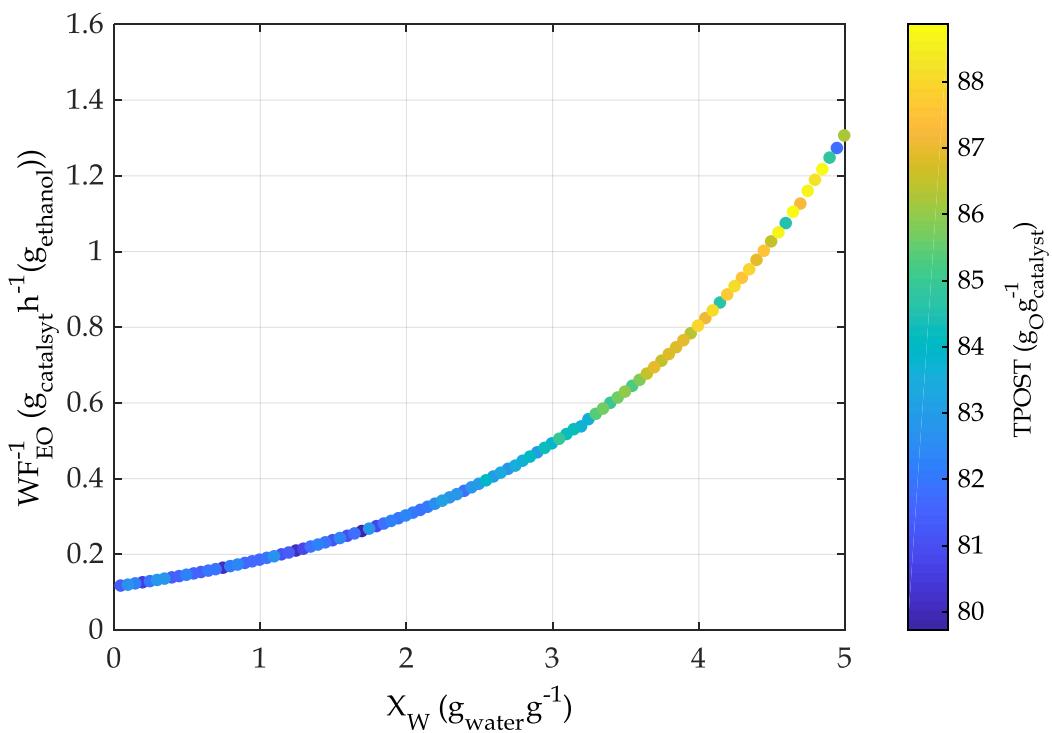


Figure 78. Comparison between the best optimisation results obtained for each X_W value.

6.3.2. Effect of the catalyst regeneration cost

The optimisation procedure has been carried out with the objective of maximising the total production and of extending the catalyst lifespan to the maximum because the catalyst is one of the main costs of the process. However, depending on the cost derived from each regeneration phase of the catalyst, the optimal operational conditions must be modified. Therefore, an optimisation procedure has been performed following the same grid search strategy to study the operating variables and changing the catalyst regeneration costs. The following cost function (I_C) has been defined to be maximised:

$$\max(I_C) = \max_{T, X_W, WF_{EO}^{-1}} n_C \cdot TPOST - n_C \cdot \exp(P_i) \quad (32)$$

The TPOST is calculated using the same estimation previously used (Equation (19)). The parameter n_c represents the total production cycles that can be carried out in 800 hours, which has resulted to be the lifespan extension limit. Finally, for each production cycle a regeneration phase must be carried out. The penalization function of this regeneration procedure has been pondered for 14 different costs following an exponential equation. Figure 79 shows the evolution of the estimated regeneration costs for each regeneration pause. This cost will be used during the optimisation procedure and it is defined for 14 penalization indexes ($P_i = [-3:10]$).

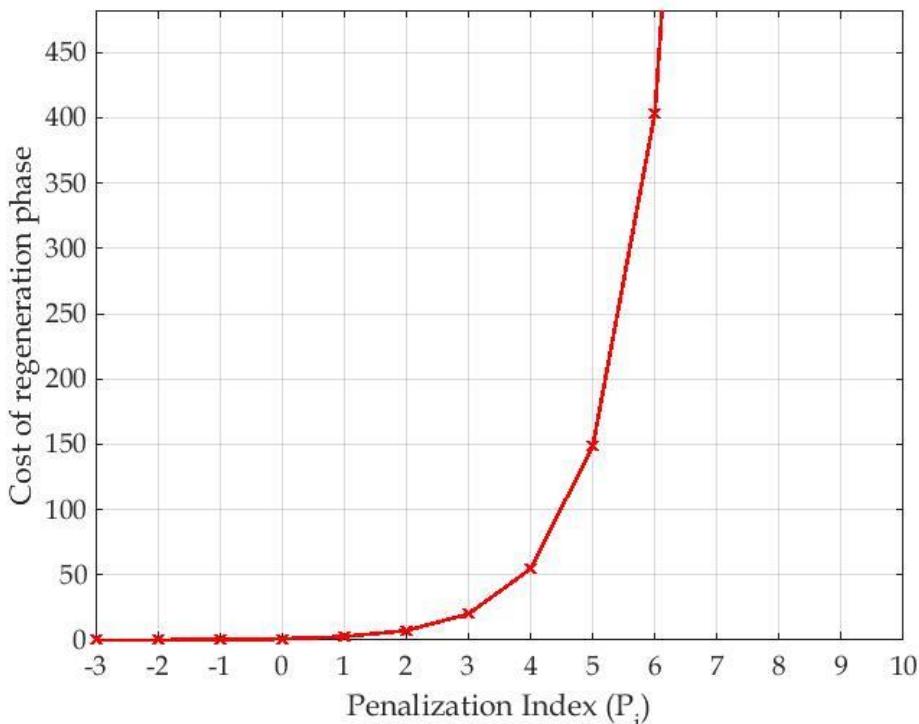


Figure 79. Cost of each regeneration procedure depending on the selected penalization index (P_i).

Therefore, for each P_i and X_W values (using the previous presented grid) a complete optimisation procedure has been carried out, optimising

simultaneously the temperature trajectories (following the f_8 profile) and the space-time.

Figure 80 and Figure 81 show that depending on the catalyst regeneration costs the optimal operational conditions move from large production steps, extending the catalyst lifespan as large as possible, to short production cycles with many production regeneration cycles when the catalyst regeneration costs decreases below a minimum. Figure 80 shows that for the penalization index 3 onwards the total production increases significantly to avoid an elevated number of costly regeneration phases. In the same way, Figure 81 shows that with the increment of the regeneration cost the number of cycles is reduced drastically.

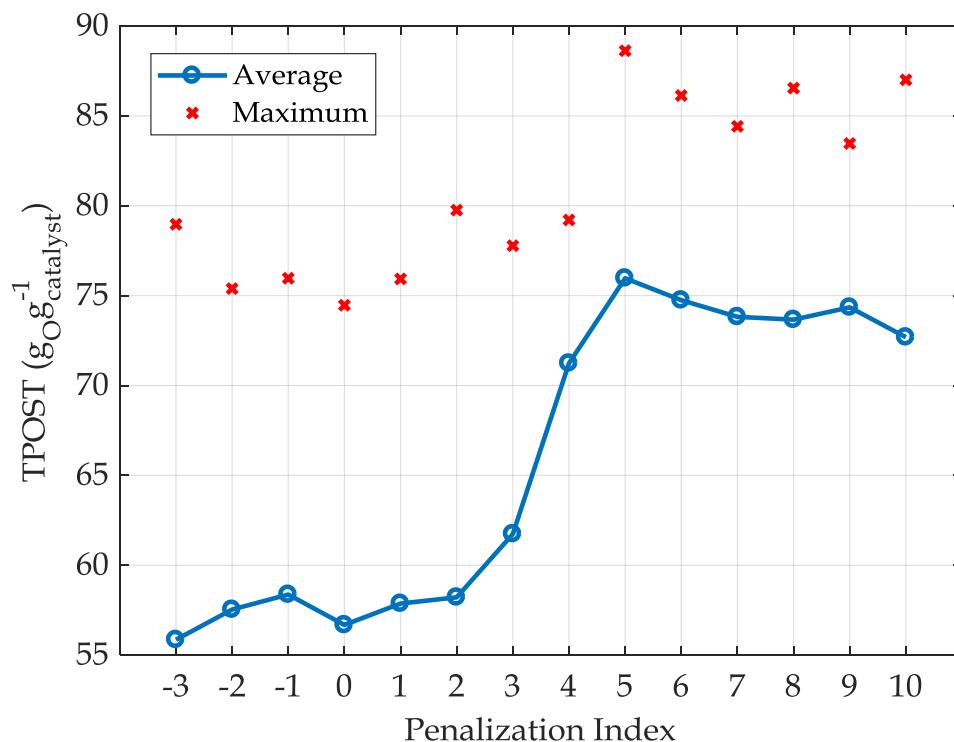


Figure 80. Average TPOST for several optimisation procedures using the proposed penalization indexes.

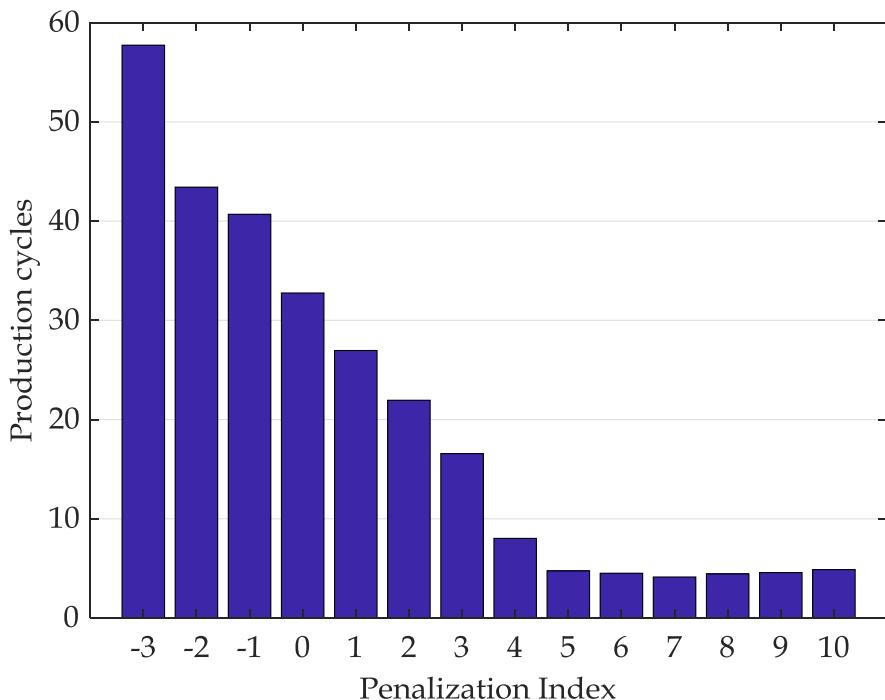


Figure 81. Total number of production-regeneration cycles depending on the penalization index.

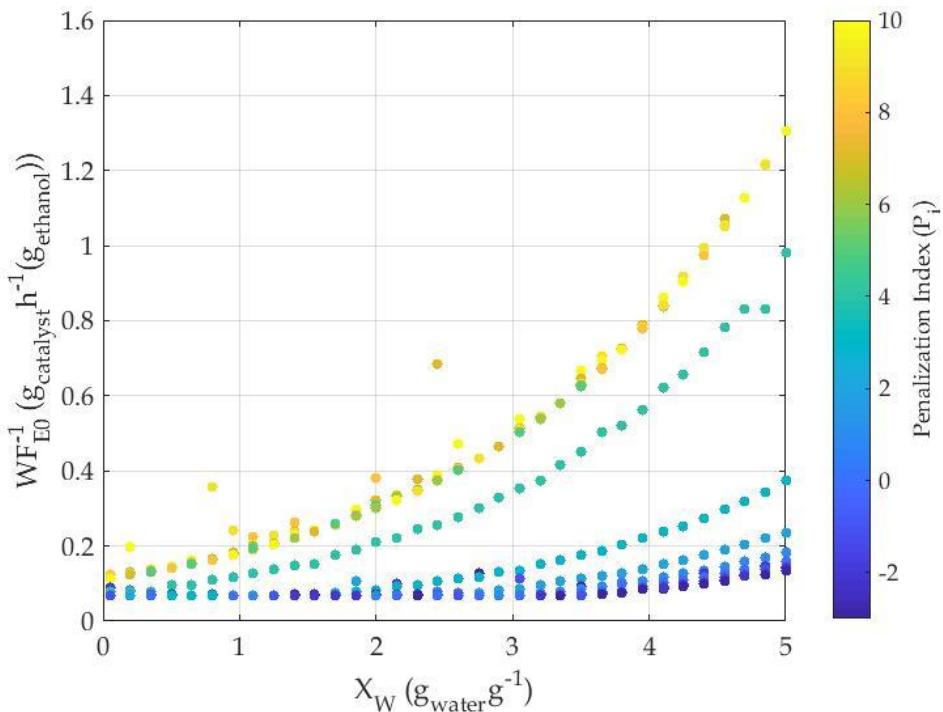


Figure 82. Best operational conditions (X_W and WF_{EO}^{-1}) for each X_W value and penalization index.

In addition, the optimal relation previously founded, between the water content in the feed and the space-time changes depending on the costs of the regeneration costs. Figure 82 shows that the relation between both variables is maintained following an exponential equation which changes in function of the penalization costs. The “colorbar” represents the penalization indexes. Notice that optimisation procedures with high P_i ($P_i > 3$) obtain the same solution curves as those obtained for only 1 production cycle (see Figure 78).

Finally, Figure 83 shows the distribution of the operating variables X_W and WF^{-1}_{EO} depending on the regeneration costs. Please note that, in this figure, the value of the optimization cost function (I_c) has been normalized in the interval [0, 1]. If the regeneration costs are too high, the operational conditions move to the maximum possible space-time and water content in order to extend to the maximum the catalyst lifespan and reduce the required number of regeneration phases. This method produces longer production steps with the maximum possible TPOST. However, if the regeneration cost is negligible, the optimal operational conditions move to short production cycles with the minimum or no water content in the feed and the minimum space-time.

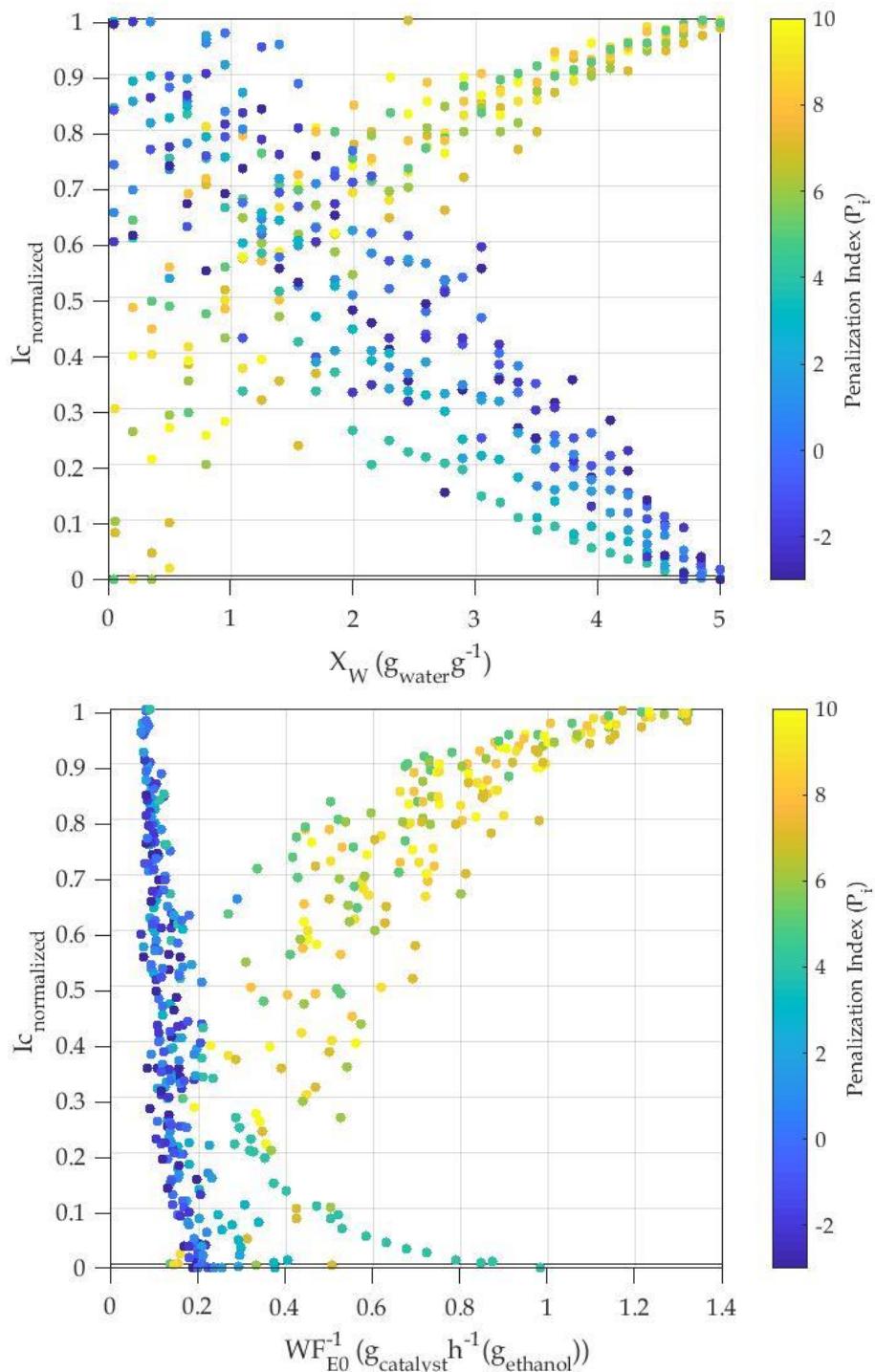


Figure 83. Results of the grid search (X_W) for each penalization index. Top: Normalized I_C results distribution for X_W . Bottom: Normalized I_C results distribution for WF_{EO}^{-1} .

6.4. Conclusions

The optimisation of the BTO process has been carried out following the defined simulation based methodology. Two optimisation procedures have been launched simultaneously. The first one has been carried out using the mechanistic model to estimate the behaviour of the process. And the second one with a significantly less computational cost (up to 50 times lower), has been carried out using the previously validated neural model of the BTO process.

The conclusions achieved by both procedures have been equivalents. On the one hand, it has been demonstrated that temperature trajectories inverse to the catalyst deactivation, where the temperature slowly increases, maximise the production of olefins per space-time. These temperature profiles are able to reach olefin's productions of almost $90 \text{ g}_\text{o}_\text{g}^{-1}\text{catalyst}$ per production cycle. On the other hand, using dynamic operational conditions for all the operating variables (temperature, space-time and water content in the feed) and not only for the temperature control, increases even more the total production, reaching TPOST of $101.72 \text{ g}_\text{o}_\text{g}^{-1}\text{catalyst}$ per production cycle.

Statistical analyses have confirmed that there exist significant differences between the proposed operational strategies. These results obtained using the dynamic trajectories of all operational variables are significantly better than those obtained with fixed temperature profiles or constant set points. The artificial neural networks have been able to dynamically calculate the optimum trajectories.

Additionally, further analyses of the operational conditions have been performed. Two aspects have been studied using a grid search methodology, the catalyst lifespan extension and the evolution of the best operational conditions depending on the catalyst regeneration costs. It has been confirmed the previous conjecture that the water content in the feed attenuate the catalyst deactivation extending the catalyst lifespan and

therefore the extension of the production cycles. Moreover, an exponential relation between the optimal values of water in the feed and the space-time has been observed. This exponential relation is maintained with changes although the regeneration costs changed.

It can be concluded that with elevated costs of the catalyst regeneration phase, the optimal operational conditions are those capable of extending the catalyst lifespan, obtaining large production cycles. In these cases, the temperature is maintained in intermediate values which slowly rise during the process in order to increase the production of olefins. The water content rises to the maximum possible in order to attenuate the catalyst deactivation and the space-time, following the exponential relation, also rises to the maximum value.

In contrast, with negligible regeneration costs, the temperature is maintained with similar values during all the process in order to maximise the conversion rate of the lump of olefins. However, the water content and the space-time decrease to the minimum in order to obtain short production cycles still maintaining an acceptable production of olefins.

CONCLUSIONS AND FUTURE WORK

This chapter will exposed the main conclusions and the future work derived from the presented work. The main conclusion is that a new methodology based in soft-computing techniques has been able to identify the most suitable operational conditions of the Bioethanol-To-Olefins (BTO) chemical process. This methodology can be easily extrapolated to other biorefinery or chemical processes.

The improvement and optimisation of complex processes, such as the BTO chemical process, can be successfully performed using soft-computing based methodology for the knowledge extraction and in the search of the most suitable operational conditions in order to maximise the selected optimisation objectives.

7.1. Conclusions and contributions

7.1.1. Modelling of the BTO process

Soft-computing modelling techniques such as Support Vector Machines and Artificial Neural Networks have been used for the modelling of the BTO process. The training dataset has been composed by experimental data of the process obtained in a laboratory scale reactor. Moreover, during the modelling procedure, additional simulated data have been included in the training datasets, as synthetic experimental data, in order to provide the models with additional information of the process dynamics.

Two modelling approaches have been explored. The soft-modelling of the process and a hybrid model structure that combines soft-modelling techniques with knowledge-based models.

The hybrid neural network based modelling approach has been able to successfully identify and reproduce the dynamics of the process. The model structure identified as HNN 2, combines the catalyst deactivation knowledge model with a neural model of the main kinetic reactions to calculate the olefins conversion rate (X_O).

This model has been compared with a well-known mechanistic model of the process, which has been used as a contrast method. The Two One-Sided Test equivalence test has been used for the model validation. According to this test, the estimations of the fitted neural model are indistinguishable from both the mechanistic model and also from the experimental data. The discrepancy of the neural model with the mechanistic model is of 0.1259 percentage points; and the coefficient of determination (R^2) for the test dataset of 0.9918.

7.1.2. Optimisation of the operational conditions

Three different evolutionary optimization strategies have been carried out in order to maximize the total production of olefins at the reactor output, while extending the catalyst lifespan. All of them improve the production of olefins when comparing with the optimisation of the process at zero time on stream (the commonly used optimisation strategy).

The optimisation with constant set-points improves the production of olefins in a 11.15 %; the optimisation using fixed-shape temperature trajectories in a 35.75 %; and the operation using dynamically generated trajectories of all the operating variables (temperature, water content and space-time) generated by the optimised ANN improves the production results in a 53.48 %.

On the one hand, the use of dynamic temperature profiles instead of constant values, significantly improves the process total production. Several temperature profile functions have been proposed, being the linear profiles and the trajectories inverse to the catalyst deactivation the best options to

maximize the total production of olefins, achieving productions 21.97 % higher than maintaining constant set-points (current operation strategy).

On the other one, Artificial Neural Networks (ANN) trained using an evolutionary algorithm have been able to successfully generate the best solutions for the main control trajectories of the BTO process, by defining their most convenient curve form and shape defining parameters. This optimisation strategy improves the total production up to 37.91 % of higher production than the reference operational strategy (constant set-points), surpassing the use of fixed temperature profiles in 15.94 percentage points.

According to the Friedman and Wilcoxon tests, there exist significant differences between the solutions provided by the ANN and those using fixed temperature profiles, specially comparing with constant operational conditions. Some fixed temperature profiles, as trajectory functions inverse to the catalyst deactivation, obtain similar results comparing with ANN generated profiles. In this point, it can be stated that the dynamic optimisation of the temperature profiles is enough to obtain near optimum solutions for the present optimisation problem with similar performance as optimising the three operating variables. However, a deeper analysis has confirmed the existence of significant differences between solutions provided by ANN and those that only use optimised temperature trajectories.

Finally, the study of the operational conditions exposed that depending on the catalyst regeneration cost the optimised solutions move to different operational conditions. With high costs the solutions try to extend the catalyst lifespan, with high quantity of water in feed and high values of space times and maintaining intermediate or slow increments of the temperature. These conditions result in long production cycles and a minimum number of regeneration phases. In contrast, when the catalyst regeneration costs are negligible, the best solutions appear as a combination of rapid production cycles with a considerable number of production cycles.

In these cases, both the best values of water content and space-time get changed maintaining an exponential relationship.

7.1.3. Actual contributions

The main contributions of the presented work are summarized below:

- An empirical model based on Artificial Neural Networks of the BTO process has been presented. A hybrid approach, including a knowledge-based model of the deactivation of the catalyst, has demonstrated to be equivalent in terms of RMSE (variability) to the experimental error of the data.
- Optimal operational strategies that maximise the production of the BTO process while extending the catalyst lifespan have been presented. During the optimisation procedure, the effect of the catalyst deactivation has been taken into account due to the relevant effect that the deactivation has in the process performance and thus in the optimal operation strategy.
- A study of the effect that the catalyst regeneration cost has in the operational conditions has been carried out. The evolution of the operational conditions as a function of this cost has been presented.

Finally, the presented contributions can be bundled in a methodology based on soft computing techniques for the study of optimal operational conditions for complex and unknown processes in general and particularly, for biomass catalytic transformation processes.

7.2. Future work and research lines

From the presented work, some future improvement possibilities have been identified specially for the optimisation procedures. The use of straight line programs (Alonso et al., 2009) or similar automatic modelling techniques could be used to search for the best trajectories or the introduction of a more complete fitness function to take into account the operational cost, the time to regenerate the catalyst, the energy cost, etc.

Additionally, the current optimisation procedure takes into account the catalyst reversible deactivation. However, depending on the process being optimised, some catalyst also presents an irreversible deactivation after each regeneration phase. Depending on the operational conditions both reversible and irreversible deactivation rates will change in a different way. Therefore, a new optimisation research line, that considers both deactivation types, should be analysed.

Finally, future developments will be necessary to develop robust controllers capable of faithfully following the dynamically optimised variable trajectories. A possible new research line could be the integration of the presented optimisation methodology in an advanced control strategy which includes an online optimisation of the operational conditions and the online model update.

BIBLIOGRAPHY

- Adib, H. et al., 2013. Support Vector Machine based modeling of an industrial natural gas sweetening plant. *Journal of Natural Gas Science and Engineering*, 14, pp.121–131.
- Ahmadi, M.A. & Pournik, M., 2016. A predictive model of chemical flooding for enhanced oil recovery purposes: Application of least square support vector machine. *Petroleum*, 2(2), pp.177–182.
- Al-Asheh, S., Mjalli, F.S. & Alfadala, H.E., 2007. Forecasting Influent-Effluent Wastewater Treatment Plant Using Time Series Analysis and Artificial Neural Network Techniques. *Chemical Product and Process Modeling*, 2(3), p.2007.
- Alonso, A., 2008. *Catalizadores alternativos y modelado cinético del proceso BTO (Bioetanol a Olefinas)*.
- Angira, R. & Babu, B.V., 2006. Optimization of process synthesis and design problems: A modified differential evolution approach. *Chemical Engineering Science*, 61(14), pp.4707–4721.
- Atias, J. a. & de Lasa, H., 2004. Adsorption and catalytic reaction in FCC catalysts using a novel fluidized CREC riser simulator. *Chemical Engineering Science*, 59(22–23), pp.5663–5669.
- Azarpour, A. et al., 2015. Catalytic activity evaluation of industrial Pd/C catalyst via gray-box dynamic modeling and simulation of hydropurification reactor. *Applied Catalysis A: General*, 489, pp.262–271.
- Babu, B.V. & Angira, R., 2006. Modified differential evolution (MDE) for optimization of non-linear chemical processes. *Computers & Chemical Engineering*, 30(6–7), pp.989–1002.
- Balabin, R.M. & Lomakina, E.I., 2011. Support vector machine regression (SVR/LS-SVM)—an alternative to neural networks (ANN) for analytical

- chemistry? Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *The Royal Society of Chemistry: Analyst*, 136(8), pp.1703–1712.
- Bansal, S., Roy, S. & Larachi, F., 2012. Support vector regression models for trickle bed reactors. *Chemical Engineering Journal*, 207–208, pp.822–831.
- Basak, D., Pal, S. & Patranabis, D.C., 2007. Support Vector Regression. *Neural Information Processing - Letters and Reviews*, 11(10), pp.203–224.
- Bhat, N. & McAvoy, T.J., 1990. Use of neural nets for dynamic modeling and control of chemical process systems. *Computers & Chemical Engineering systems*, 14(May), pp.573–582.
- Bhat, N.V. et al., 1990. Modeling chemical process systems via neural computation. *IEEE Control Systems Magazine*, 10(3), pp.24–30.
- Bi, J. et al., 2011. C₂-C₄ light olefins from bioethanol catalyzed by Ce-modified nanocrystalline HZSM-5 zeolite catalysts. *Applied Catalysis B: Environmental*, 107(1–2), pp.68–76.
- Braik, M., Sheta, A. & Arieqat, A., 2008. A Comparison between GAs and PSO in Training ANN to Model the TE Chemical Process Reactor. In *Proceedings of the AISB 2008 Symposium on Swarm Intelligence Algorithms and Applications*. pp. 24–31.
- Brierley, P.D., 1998. *Some Practical Applications of Neural Networks in the Electricity Industry*. Cranfield University.
- Chang, C.-C. & Lin, C.J., 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2, pp.1–27.
- Chen, L. et al., 2004. Modelling and optimization of fed-batch fermentation processes using dynamic neural networks and genetic algorithms. *Biochemical Engineering Journal*, 22(1), pp.51–61.
- Chitralekha, S.B. & Shah, S.L., 2010. Application of support vector regression for developing soft sensors for nonlinear processes. *Canadian Journal of Chemical Engineering*, 88(5), pp.696–709.

- Corma, A. et al., 2005. Integrating high-throughput characterization into combinatorial heterogeneous catalysis: unsupervised construction of quantitative structure/property relationship models. *Journal of Catalysis*, 232(2), pp.335–341.
- Cunill, F., Iborra, M. & Tejero, J., 2010. *Reactores químicos* U. D. E. Barcelona, ed., Barcelona, Universitat D E.
- Cybenko, G., 1989. Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control , Signals , and Systems*, 2, pp.303–314.
- Dach, J. et al., 2016. The use of neural modelling to estimate the methane production from slurry fermentation processes. *Renewable and Sustainable Energy Reviews*, 56, pp.603–610.
- David F. Bauer, 1972. Constructing Confidence Sets Using Rank Statistics. *Journal of the American Statistical Association*, 67(339), pp.687–690.
- Demšar, J., 2006. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, 7, pp.1–30.
- Demuth, H., Beale, M. & Hagan, M., 2010. *Neural Network Toolbox™ 6 User's Guide* 2010th ed., The MathWorks.
- Dossumov, K. et al., 2012. Production of olefins from bioethanol. Catalysts, mechanism. *Chemical Bulletin of Kazakh National University*, 4(68), pp.42–49.
- Dragoi, E.-N. et al., 2013. Optimization methodology based on neural networks and self-adaptive differential evolution algorithm applied to an aerobic fermentation process. *Applied Soft Computing*, 13(1), pp.222–238.
- El-Atta, A.H.A. & Hassanien, A.E., 2017. Two-class support vector machine with new kernel function based on paths of features for predicting chemical activity. *Information Sciences*, 403–404, pp.42–54.
- Farshad, F. et al., 2011. Separation of toluene/n-heptane mixtures experimental, modeling and optimization. *Chemical Engineering Journal*, 173(1), pp.11–18.
- Festel, G. et al., 2014. Modelling production cost scenarios for biofuels and

fossil fuels in Europe. *Journal of Cleaner Production*, 66, pp.242–253.

Fissore, D., Barresi, A. a. & Manca, D., 2004. Modelling of methanol synthesis in a network of forced unsteady-state ring reactors by artificial neural networks for control purposes. *Chemical Engineering Science*, 59(19), pp.4033–4041.

Food and Agricultural Policy Research Institute, 2011. *FAPRI-ISU 2011 World Agricultural Outlook*,

Galvan, I.M. et al., 1996. THE USE OF NEURAL NETWORKS FOR FITTING COMPLEX KINETIC DATA. *Computers & Chemical Engineering*, 20(12), pp.1451–1465.

Garcia, S. & Herrera, F., 2008. An Extension on “Statistical Comparisons of Classifiers over Multiple Data Sets” for all Pairwise Comparisons. *Journal of machine learning research*, 9, pp.2677–2694.

Gayubo, A.G. et al., 2012. Deactivation Kinetics of a HZSM-5 Zeolite Catalyst Treated with Alkali for the Transformation of Bio-Ethanol into Hydrocarbons. *American Institute of Chemical Engineers*, 58(2), pp.526–537.

Gayubo, A.G., Alonso, A., Valle, B., Aguayo, A.T., et al., 2010. Kinetic Model for the Transformation of Bioethanol into Olefins over a HZSM-5 Zeolite Treated with Alkali. *Industrial & Engineering Chemistry Research*, 49(21), pp.10836–10844.

Gayubo, A.G. et al., 2011. Kinetic modelling for the transformation of bioethanol into olefins on a hydrothermally stable Ni-HZSM-5 catalyst considering the deactivation by coke. *Chemical Engineering Journal*, 167(1), pp.262–277.

Gayubo, A.G., Alonso, A., Valle, B., Aguayo, A.T., et al., 2010. Selective production of olefins from bioethanol on HZSM-5 zeolite catalysts treated with NaOH. *Applied Catalysis B: Environmental*, 97(1–2), pp.299–306.

Ghadirian, A. & Zekri, M., 2011. MIMO Nonlinear Dynamic Systems Identification Using Fully Recurrent Wavelet Neural Network. In *2nd International Conference on Control, Instrumentation and Automation*

- (ICCIA). pp. 1113–1118.
- Ghaedi, M. et al., 2014. Least square-support vector (LS-SVM) method for modeling of methylene blue dye adsorption using copper oxide loaded on activated carbon: Kinetic and isotherm study. *Journal of Industrial and Engineering Chemistry*, 20(4), pp.1641–1649.
- Ghahraloud, H., Farsi, M. & Rahimpour, M.R., 2017. Modeling and optimization of an industrial Claus process: Thermal and catalytic section. *Journal of the Taiwan Institute of Chemical Engineers*, 76, pp.1–9.
- Goldberd, D.E., 1989. *Genetic algorithms in search, optimization, and machine learning* 1st ed., Addison-Esley Longman Publishing Co.
- Gonzaga, J.C.B. et al., 2009. ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Computers and Chemical Engineering*, 33(1), pp.43–49.
- Gueguim Kana, E.B. et al., 2012. Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm. *Renewable Energy*, 46, pp.276–281.
- Happel, J., Sellers, P.H. & Otarad, M., 1990. Mechanistic study of chemical reaction systems. *Industrial & Engineering Chemistry Research*, 29, pp.1057–1064.
- Holeňa, M. & Baerns, M., 2003. Feedforward neural networks in catalysis: A tool for the approximation of the dependency of yield on catalyst composition, and for knowledge extraction. *Catalysis Today*, 81(3), pp.485–494.
- Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems* U. of M. Press, ed.,
- Hong, W.C., 2011. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. *Energy*, 36(9), pp.5568–5578.
- Hornik, K., Stinchcombe, M. & White, H., 1989. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), pp.359–366.

- Hornik, K., Stinchcombe, M. & White, H., 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3(5), pp.551–560.
- Hosen, M.A., Hussain, M.A. & Mjalli, F.S., 2011. Control of polystyrene batch reactors using neural network based model predictive control (NNMPC): An experimental investigation. *Control Engineering Practice*, 19(5), pp.454–467.
- Hough, B.R. et al., 2017. Application of machine learning to pyrolysis reaction networks: Reducing model solution time to enable process optimization. *Computers and Chemical Engineering*, 104, pp.56–63.
- Hu, Q. et al., 2016. Short-Term Wind Speed Forecasting Based on Heteroscedastic Support Vector Regression. *IEEE Transaction on Sustainable Energy*, 7(1), pp.241–249.
- Huber, G.W. & Corma, A., 2007. Synergies between bio- and oil refineries for the production of fuels from biomass. *Angewandte Chemie (International ed. in English)*, 46(38), pp.7184–7201.
- Inaba, M. et al., 2007. Production of olefins from ethanol by Fe-supported zeolite catalysts. *Journal of Chemical Technology and Biotechnology*, (9), pp.638–646.
- International Energy Agency, 2011. *Technology Roadmap: Biofuels for Transport*,
- Jagannathant, S. & Lewis, F.L., 1996. Identification of Nonlinear Dynamical Systems Using Multilayered Neural Networks *. *Automatica*, 32(12), pp.1707–1712.
- Jin, H. et al., 2015. Multi-model adaptive soft sensor modeling method using local learning and online support vector regression for nonlinear time-variant batch processes. *Chemical Engineering Science*, 131, pp.282–303.
- Jin, Y., 2011. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, 1(2), pp.61–70.
- John M. Chambers, 1983. *Graphical methods for data analysis* Wadsworth

- International Group, ed.,
- Jong, E. De et al., 2012. Product developments in the bio-based chemicals arena. *Biofuels, Bioproducts and Biorefining*, (6), pp.606–624.
- Jong, E. De & Jungmeier, G., 2015. Biorefinery Concepts in Comparison to Petrochemical Refineries. In *Industrial Biorefineries and White Biotechnology*. pp. 3–33.
- Kajero, O.T. et al., 2017. Meta-modelling in chemical process system engineering. *Journal of the Taiwan Institute of Chemical Engineers*, 73, pp.135–145.
- Kashani, M.N. & Shahhosseini, S., 2010. A methodology for modeling batch reactors using generalized dynamic neural networks. *Chemical Engineering Journal*, 159(1–3), pp.195–202.
- Kavousi-Fard, A., Samet, H. & Marzbani, F., 2014. A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting. *Expert Systems with Applications*, 41(13), pp.6047–6056.
- Kazempour, H., Pourfayaz, F. & Mehrpooya, M., 2017. Modeling and multi-optimization of thermal section of Claus process based on kinetic model. *Journal of Natural Gas Science and Engineering*, 38, pp.235–244.
- Ko, M., Tiwari, A. & Mehnen, J., 2010. A review of soft computing applications in supply chain management. *Applied Soft Computing*, 10(3), pp.661–674.
- Koçar, G. & Civa, N., 2013. An overview of biofuels from energy crops : Current status and future prospects. , 28, pp.900–916.
- Koller, R.W. & Ricardez-sandoval, L.A., 2017. A Dynamic Optimization Framework for Integration of Design , Control and Scheduling of Multi-product Chemical Processes under Disturbance and Uncertainty. *Computers and Chemical Engineering*.
- Kordabadi, H. & Jahanmiri, a., 2007. A pseudo-dynamic optimization of a dual-stage methanol synthesis reactor in the face of catalyst deactivation. *Chemical Engineering and Processing: Process Intensification*, 46(12), pp.1299–1309.

- Kordabadi, H. & Jahanmiri, a., 2005. Optimization of methanol synthesis reactor using genetic algorithms. *Chemical Engineering Journal*, 108(3), pp.249–255.
- Laguna, M. & Martí, R., 2002. Neural network prediction in a system for optimizing simulations. *IIE Transactions*, 34, pp.273–282.
- Landolina, S. & Maltsoglou, I., 2017. How2Guide for Bioenergy Roadmap Development and Implementation. *Iea*.
- Ławryńczuk, M., 2011. On-line set-point optimisation and predictive control using neural Hammerstein models. *Chemical Engineering Journal*, 166(1), pp.269–287.
- Lee, D.E. et al., 2005. Weighted Support Vector Machine for Quality Estimation in the Polymerization Process. *Ind. Eng. Chem. Res.*, 44(7), pp.2101–2105.
- Lightbody, G. et al., 1994. Neural network modelling of a polymerisation reactor. In *CONTROL*. pp. 237–242.
- López, M.E. et al., 2017. Modelling the removal of volatile pollutants under transient conditions in a two-stage bioreactor using artificial neural networks. *Journal of Hazardous Materials*, 324, pp.100–109.
- Løvik, I., Hillestad, M. & Hertzberg, T., 1998. Long Term Dynamic Optimization of a Catalytic Reactor System. *Computers & Chemical Engineering*, 22(98), pp.6–9.
- Løvik, I., Hillestad, M. & Hertzberg, T., 1999. Modeling and optimization of a reactor system with deactivating catalyst. *Computers & Chemical Engineering*, 23, pp.S839–S842.
- Marin, G.B. & Yablonsky, G.S., 2011. *Kinetics of Chemical Reactions: Decoding Complexity*,
- Marulanda, J.F., Barco, M.G. & López, J.A., 2007. *Control Inteligente De Un Reactor Químico*, Cali.
- Meert, K. & Rijckaert, M., 1998. Intelligent modelling in the chemical process industry with neural networks: a case study. *Computers & Chemical*

- Engineering*, 22(98), pp.S587–S593.
- Mikulandrić, R. et al., 2014. Artificial neural network modelling approach for a biomass gasification process in fixed bed gasifiers. *Energy Conversion and Management*, 87, pp.1210–1223.
- Mitra, K., Deb, K. & Gupta, S.K., 1998. Multiobjective Dynamic Optimization of an Industrial Nylon 6 Semibatch Reactor Using Genetic Algorithm. *Journal of Applied Polymer Science*, 69, pp.69–87.
- Molga, E.J., 2003. Neural network approach to support modelling of chemical reactors: problems, resolutions, criteria of application. *Chemical Engineering and Processing: Process Intensification*, 42(8–9), pp.675–695.
- Mosier, N. et al., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource technology*, 96(6), pp.673–86.
- Mujtaba, I.M., Aziz, N. & Hussain, M. a., 2006. Neural Network Based Modelling and Control in Batch Reactor. *Chemical Engineering Research and Design*, 84(8), pp.635–644.
- Muñoz López, C.A. et al., 2018. A process simulator interface for multiobjective optimization of chemical processes. *Computers and Chemical Engineering*, 109, pp.119–137.
- Myles Hollander, Douglas A. Wolfe & Eric Chicken, 2014. *Nonparametric Statistical Methods* 3rd Editio. Wiley, ed.,
- Nair, V. V. et al., 2016. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. *Bioresource Technology*, 217, pp.90–99.
- Narendra, K.S. & Parthasarathy, K., 1990. Identification and Control of Dynamical Systems Using Neural Networks. *IEEE Transactions on Neural Networks*, I(I).
- Nascimento, C.A.O., Giudici, R. & Guardani, R., 2000. Neural network based approach for optimization of industrial chemical processes. *Computers & Chemical Engineering*, 24, pp.2303–2314.
- Norgaard, M. et al., 2003. *Neural Networks for Modelling and Control of Dynamic Systems* Springer., London: Springer.

- Oar-Arteta, L. et al., 2015. Effect of Operating Conditions on Dimethyl Ether Steam Reforming over a CuFe₂O₄/γ-Al₂O₃ Bifunctional Catalyst. *Industrial & Engineering Chemistry Research*, 54, pp.9722–9732.
- Ochoa, S., Repke, J.-U. & Wozny, G., 2010. Integrating real-time optimization and control for optimal operation: Application to the bio-ethanol process. *Biochemical Engineering Journal*, 53(1), pp.18–25.
- Oduguwa, V., Tiwari, a. & Roy, R., 2005. Evolutionary computing in manufacturing industry: an overview of recent applications. *Applied Soft Computing*, 5(3), pp.281–299.
- Ong, Y.S. et al., 2004. Surrogate-assisted evolutionary optimization frameworks for high-fidelity engineering design problems. In *Knowledge Incorporation in Evolutionary Computation*. pp. 307–331.
- Papadokonstantakis, S. et al., 2005. Variable selection and data pre-processing in NN modelling of complex chemical processes. *Computers & Chemical Engineering*, 29(7), pp.1647–1659.
- Pekar, M., 2004. What can kinetics learn from rational thermodynamics. *Chemical Engineering Science*, 59(19), pp.4103–4112.
- Pérez-Uriarte, P. et al., 2017. Deactivation kinetics for the conversion of dimethyl ether to olefins over a HZSM-5 zeolite catalyst. *Chemical Engineering Journal*, 311, pp.367–377.
- Pérez-Uriarte, P., Ateka, A., Gamero, M., et al., 2016. Effect of the Operating Conditions in the Transformation of DME to olefins over a HZSM-5 Zeolite Catalyst. *Industrial and Engineering Chemistry Research*, 55(23), pp.6569–6578.
- Pérez-Uriarte, P., Ateka, A., Aguayo, A.T., et al., 2016. Kinetic model for the reaction of DME to olefins over a HZSM-5 zeolite catalyst. *Chemical Engineering Journal*, 302, pp.801–810.
- Pollard, J.F. et al., 1992. Process identification using neural networks. *Computers & Chemical Engineering*, 16(4), p.1992.
- Rahimpour, M.R., Shayanmehr, M. & Nazari, M., 2011. Modeling and Simulation of an Industrial Ethylene Oxide (EO) Reactor Using Artificial

- Neural Networks (ANN). *Industrial & Engineering Chemistry Research*, 50(10), pp.6044–6052.
- Rajesh, J.K. et al., 2001. Multi-objective optimization of industrial hydrogen plants. *Chemical Engineering Science*, 56(3), pp.999–1010.
- Robinson, A.P., Duursma, R. a & Marshall, J.D., 2005. A regression-based equivalence test for model validation: shifting the burden of proof. *Tree physiology*, 25(7), pp.903–913.
- Rodemerck, U. et al., 2004. Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials. *Applied Surface Science*, 223(1-3), pp.168–174.
- Da Ros, S. et al., 2017. Modelling the effects of reaction temperature and flow rate on the conversion of ethanol to 1,3-butadiene. *Applied Catalysis A: General*, 530, pp.37–47.
- Sadeghassadi, M. et al., 2018. Application of neural networks for optimal-setpoint design and MPC control in biological wastewater treatment. *Computers and Chemical Engineering*, 115, pp.150–160.
- Salcedo-Sanz, S. et al., 2011. Short term wind speed prediction based on evolutionary support vector regression algorithms. *Expert Systems with Applications*, 38(4), pp.4052–4057.
- Santamaría-Bonfil, G., Reyes-Ballesteros, A. & Gershenson, C., 2016. Wind speed forecasting for wind farms: A method based on support vector regression. *Renewable Energy*, 85, pp.790–809.
- Serna, P. et al., 2008. Combining high-throughput experimentation, advanced data modeling and fundamental knowledge to develop catalysts for the epoxidation of large olefins and fatty esters. *Journal of Catalysis*, 258(1), pp.25–34.
- Serra, J.M. et al., 2007. Soft Computing Techniques Applied to Combinatorial Catalysis: A New Approach for the Discovery and Optimization of Catalytic Materials. *QSAR & Combinatorial Science*, 26(1), pp.11–26.
- Sheskin, D.J., 2007. *Handbook of Parametric and Nonparametric Statistical Procedures* 4th Editio., Chapman & Hall/CRC.

- Shi, X., Tong, C. & Wang, L., 2016. Evolutionary Optimization with Adaptive Surrogates and its Application in Crude Oil Distillation. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. pp. 1-8.
- Smola, A.J. & Schölkopf, B., 2003. A Tutorial on Support Vector Regression *. *Statistics and Computing*.
- Solgaard, A. et al., 2011. *Biofuels Vital Graphics*,
- Soltanali, S. et al., 2014. Neural network and genetic algorithm for modeling and optimization of effective parameters on synthesized ZSM-5 particle size. *Materials Letters*, 136, pp.138–140.
- Sousa, Z.S.B. et al., 2016. Ethanol conversion into olefins and aromatics over HZSM-5 zeolite: Influence of reaction conditions and surface reaction studies. *Journal of Molecular Catalysis A: Chemical*, 422, pp.266–274.
- Tian, Y., Zhang, J. & Morris, J., 2001. Modeling and Optimal Control of a Batch Polymerization Reactor Using a Hybrid Stacked Recurrent Neural Network Model. *Industrial & Engineering Chemistry Research*, 40(21), pp.4525–4535.
- Toch, K., Thybaut, J.W. & Marin, G.B., 2015. A Systematic Methodology for Kinetic Modeling of Chemical Reactions Applied to n-Hexane Hydroisomerization. *American Institute of Chemical Engineers*, 61, pp.880–892.
- Valero, S. et al., 2009. DoE framework for catalyst development based on soft computing techniques. *Computers & Chemical Engineering*, 33(1), pp.225–238.
- Velez-Langs, O., 2005. Genetic algorithms in oil industry: An overview. *Journal of Petroleum Science and Engineering*, 47(1-2), pp.15–22.
- Walker, E. & Nowacki, A.S., 2011. Understanding equivalence and noninferiority testing. *Journal of General Internal Medicine*, 26(2), pp.192–196.
- Wang, J. et al., 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Applied Energy*, 94, pp.65–70.

- Wei, M. et al., 2015. Optimization of catalytic Naphtha reforming process based on modified differential evolution algorithm. *IFAC-PapersOnLine*, 48(8), pp.373–378.
- Yang, H., Chan, L. & King, I., 2002. Support Vector Machine Regression for Volatile Stock Market Prediction. *Intelligent Data Engineering and Automated*, pp.391–396.
- Yang, S.H., Chung, P.W.H. & Ellow, B.W.B.F., 1999. Multi-Stage modelling of a semibatch polymerization reactor using artificial neural networks. *Trans IChemE*, 77, pp.779–783.
- Yee, A.K.Y., Ray, A.K. & Rangaiah, G.P., 2003. Multiobjective optimization of an industrial styrene reactor. *Computers & Chemical Engineering*, 27, pp.111–130.
- Yi-Fan, H. et al., 2017. Modeling of expanded granular sludge bed reactor using artificial neural network. *Journal of Environmental Chemical Engineering*, 5(3), pp.2142–2150.
- Yu, D.L. & Gomm, J.B., 2003. Implementation of neural network predictive control to a multivariable chemical reactor. *Control Engineering Practice*, 11(11), pp.1315–1323.
- Yüzgeç, U., 2010. Performance comparison of differential evolution techniques on optimization of feeding profile for an industrial scale baker's yeast fermentation process. *ISA transactions*, 49(1), pp.167–176.
- Yüzgeç, U., Türker, M. & Hocalar, A., 2009. On-line evolutionary optimization of an industrial fed-batch yeast fermentation process. *ISA transactions*, 48(1), pp.79–92.
- Zamarreño, J.M. & Vega, P., 1998. State space neural network. Properties and application. *Neural networks : the official journal of the International Neural Network Society*, 11(6), pp.1099–1112.

NOMENCLATURE

a: activity, Eq. (9).

E, G, O, P: ethylene, lump of gasoline (butane and C5+), lump of propylene + butenes, and lump of light paraffins (C4–), respectively.

E_i , E_w : activation energy of i step in the kinetic scheme, and parameter that relates k_w parameter to temperature, kJ/mol.

E_{di} , E_{dw} : activation energy of kinetic constant for deactivation by coke formed from i component, and parameter that relates k_{dw} parameter to temperature, kJ/mol F_{EO} mass flow rate of ethanol in the feed, gh^{-1} .

k_i , k_i^* : kinetic constant of i step in the kinetic scheme, at temperature T and at T^* reference temperature (573 K) respectively, $g_{ethanol} g^{-1} catalyst h^{-1}$.

k_{di} , k_{di}^* : kinetic constant for deactivation by coke formed from i component, at temperature T and at T^* reference temperature (573 K) respectively, h^{-1} .

k_w , k_w^* : parameter that quantifies the resistance to the formation of i component in the corresponding reaction step due to the presence of water in the reaction medium, at temperature T and at T^* reference temperature (573 K), dimensionless.

k_{dw} , k_{dw}^* : parameter that quantifies the formation of coke due to the presence of water in the reaction medium, at temperature T and at T^* reference temperature(573 K) respectively, dimensionless.

\bar{M} : average molecular weight of organic components, $kg mol^{-1}$.

m_0 : mass flow rate of organic components, $g h^{-1}$.

n_d : exponent in Eq. (10).

R: constant of gases, $\text{kJ mol}^{-1}\text{K}^{-1}$.

$(r_i)_0$, r_i : rate of formation of i component at zero time on stream and at any time on stream, $\text{g}_i \text{ component g}_{\text{ethanol}}^{-1} \text{ h}^{-1} \text{ g}_{\text{organic components}}^{-1} \text{ catalyst}^{-1}$.

T: temperature, K.

t: time on stream, h.

u: gas linear velocity, m h^{-1} .

X_i : weight fraction of i component by mass unit of organic components.

X_E , X_G , X_O , X_P : weight fraction of ethylene, lump of gasoline (butane and C5+), lump of propylene + butenes, and lump of light paraffins C4–, respectively, by mass unit of organic components.

X_W , X_{W0} , X_{Wf} : water/organic component mass ratio in the reaction medium, in the feed, and that formed, respectively.

Z: total length of the reactor, m.

ε : bulk porosity.

$\theta(X_W)$, $\theta_d(X_W)$: functions for quantifying the effect of water in the reaction medium on the reaction rates in the kinetic scheme and on the kinetics of deactivation by coke, respectively.

ρ : catalyst density, kg m^{-3} .

Z: dimensionless longitudinal coordinate of the reactor (z/Z).

ANEX A: MODELLING DATA

This annex presents the difference between the original experimental data and the synthetically generated data using the mechanistic model.

Figure A. 1 to Figure A. 3 show the operational conditions available in both datasets for a discrete catalyst activity level for each of the principal operating variables (temperature, water content in the feed and space-time). Red points represent the experimental data available for the different activity levels, while the blue ones represent the synthetic data for the same activity levels.

As can be seen, the synthetic data explore new operational areas where the constant original experimental conditions do not provide information about the process behaviour. Moreover, these new data presents dynamic operational conditions that provide new information about the process dynamics.

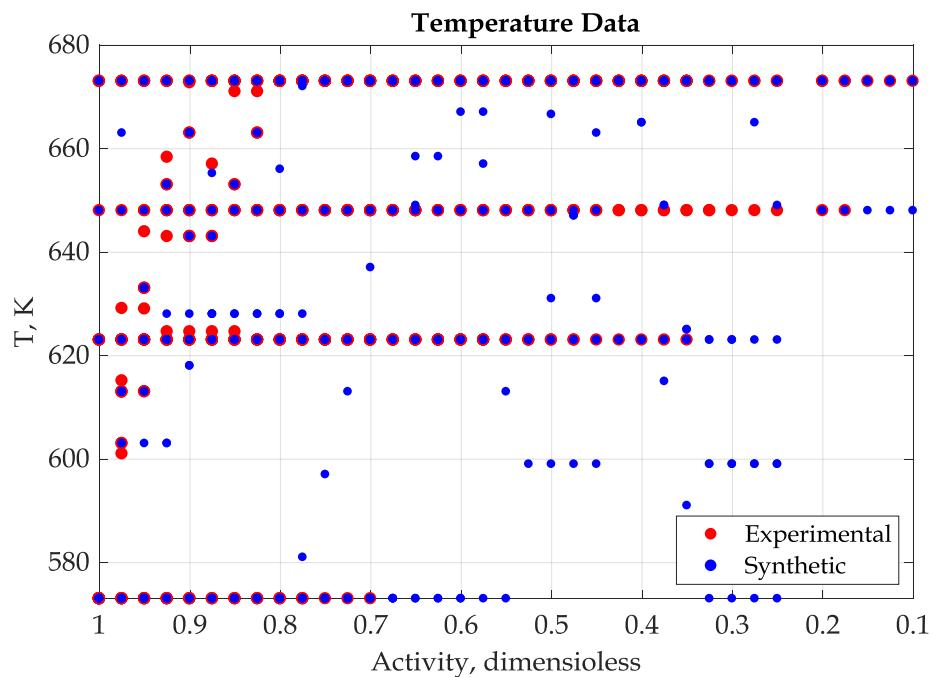


Figure A. 1. Temperature experimental (red) and synthetic (blue) data used in the modelling procedure at discrete catalyst deactivation states.

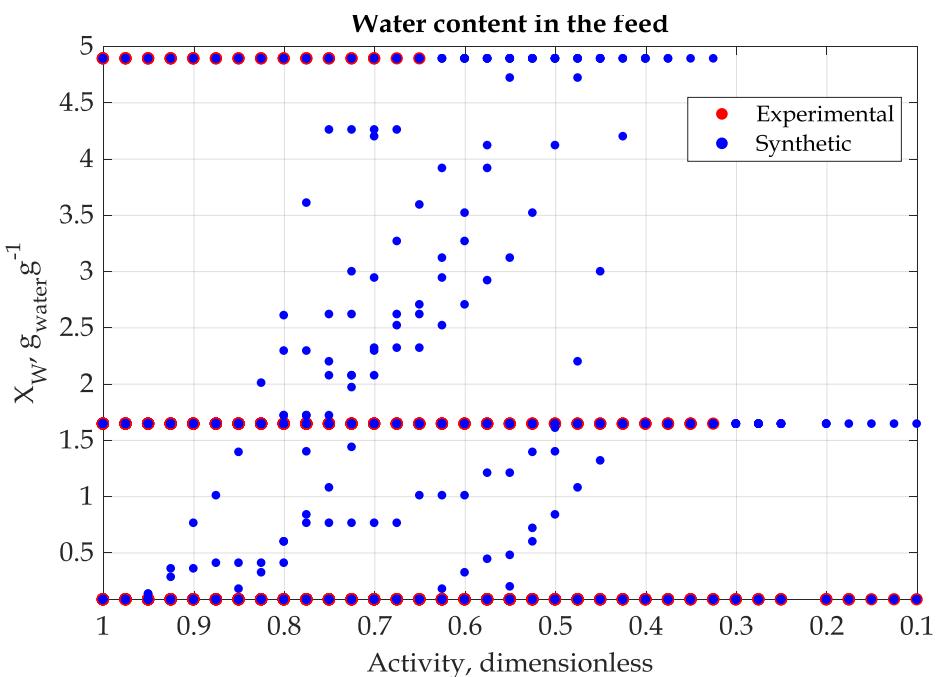


Figure A. 2. Experimental (red) and synthetic (blue) data of the water content used in the modelling procedure at discrete catalyst deactivation states.

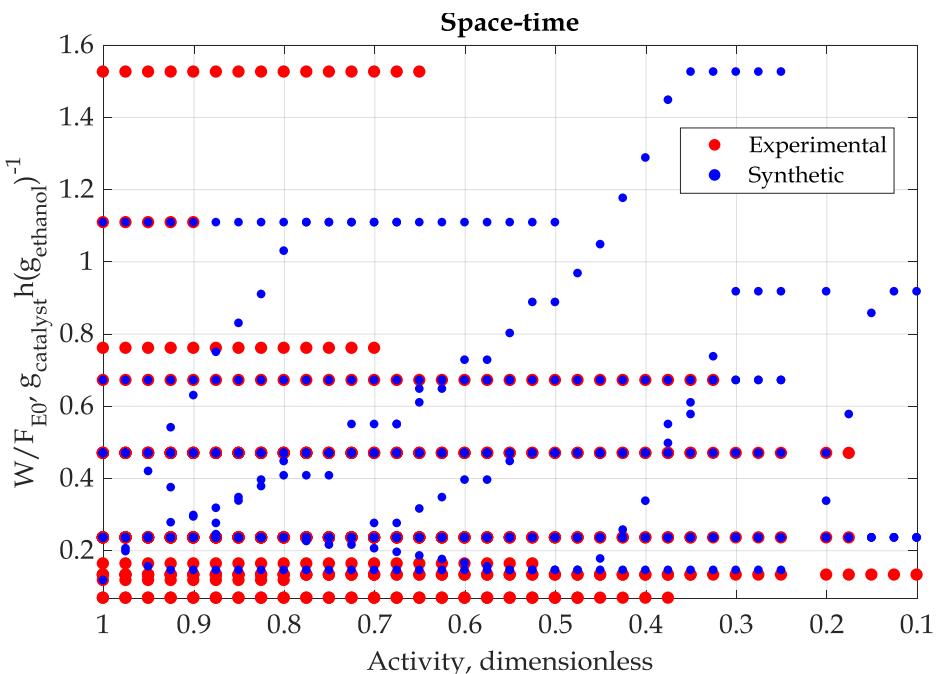


Figure A. 3. Experimental (red) and synthetic (blue) data of the space-time variable used in the modelling procedure at discrete catalyst deactivation states.

ANEX B: OPTIMISATION RESULTS

This annex presents the best results obtained in the optimisation of the operational conditions of the BTO process. The complete results obtained using each operational strategy and proposed trajectory types are shown.

Figure B. 1 to Figure B. 10 show the behaviour of the BTO process with the best operational conditions found using constant set-points (Figure B. 1) and fixed-shaped temperature trajectories. For each trajectory the best solution found using during the optimisation procedure the mechanistic model and the hybrid neural model are presented in the top and bottom figures respectively.

Finally, the best solutions found using the neuroevolution strategy in order to generate dynamic trajectories for each of the operating variables are presented. Figure B. 11 to Figure B. 15 show the behaviour of the BTO process and the operational conditions for each tested number of hidden neurons ($n = 4$ to $n = 8$).

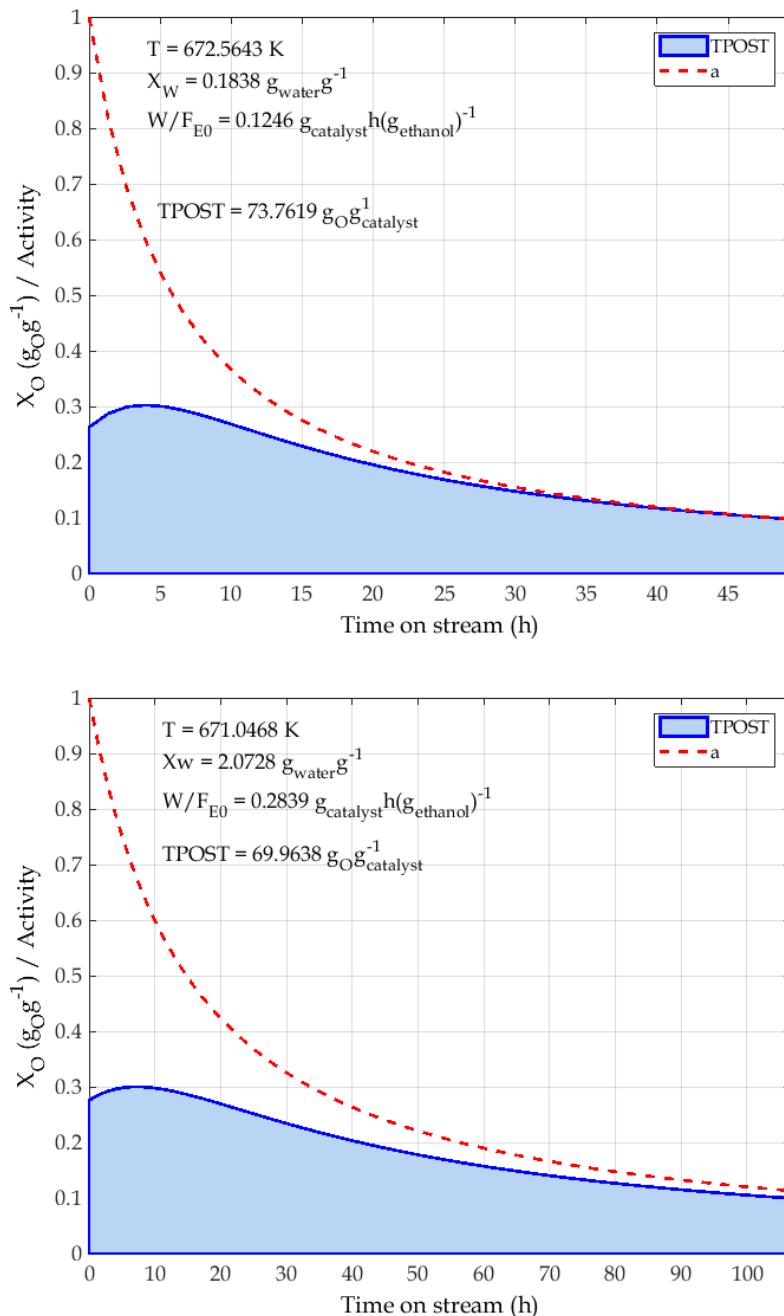


Figure B. 1. BTO process performance and operational conditions for the best solution obtained with constant set-points f_0 (Top: MECH solution; Bottom: HNN solution).

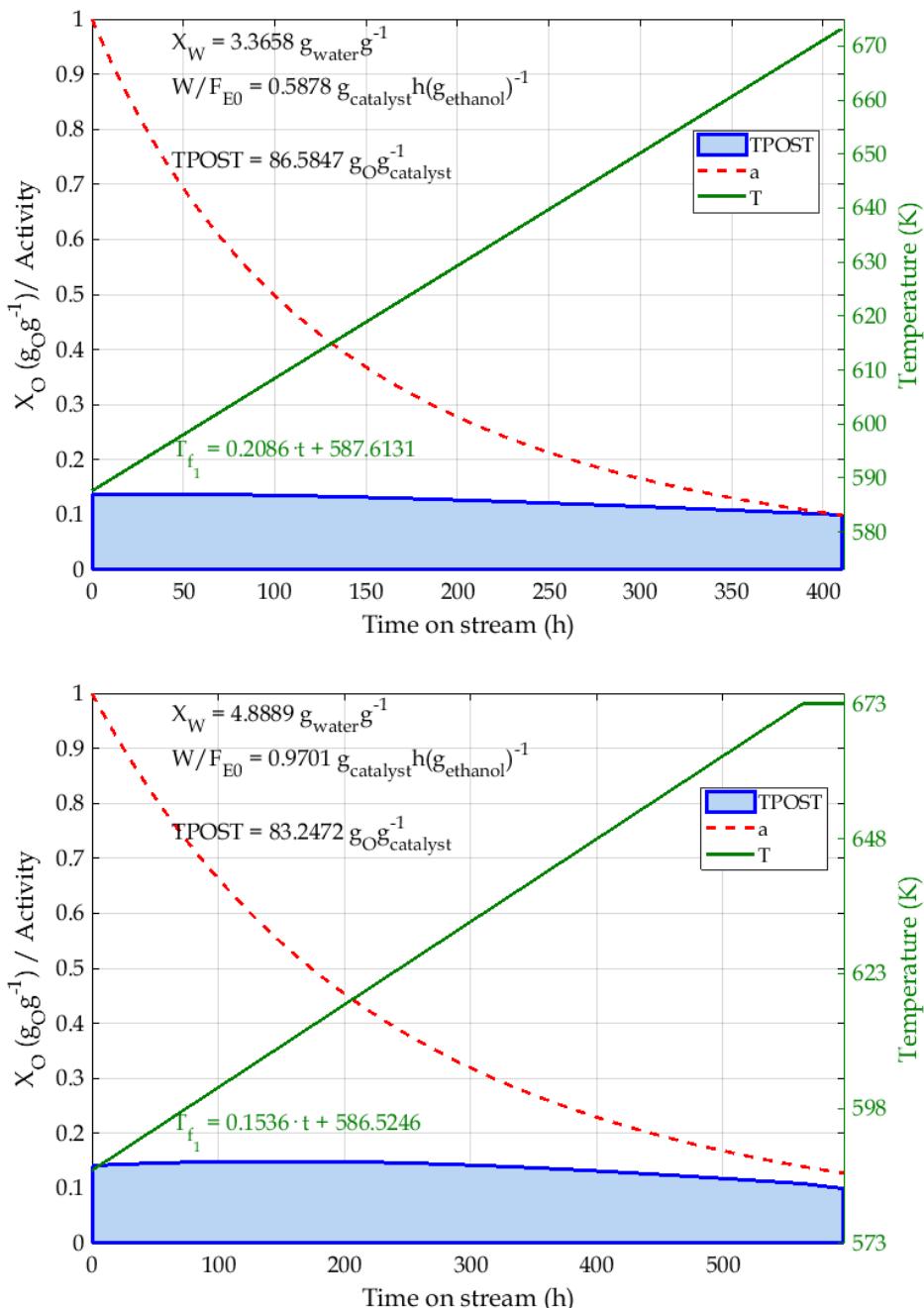


Figure B. 2. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_1 (Top: MECH solution; Bottom: HNN solution).

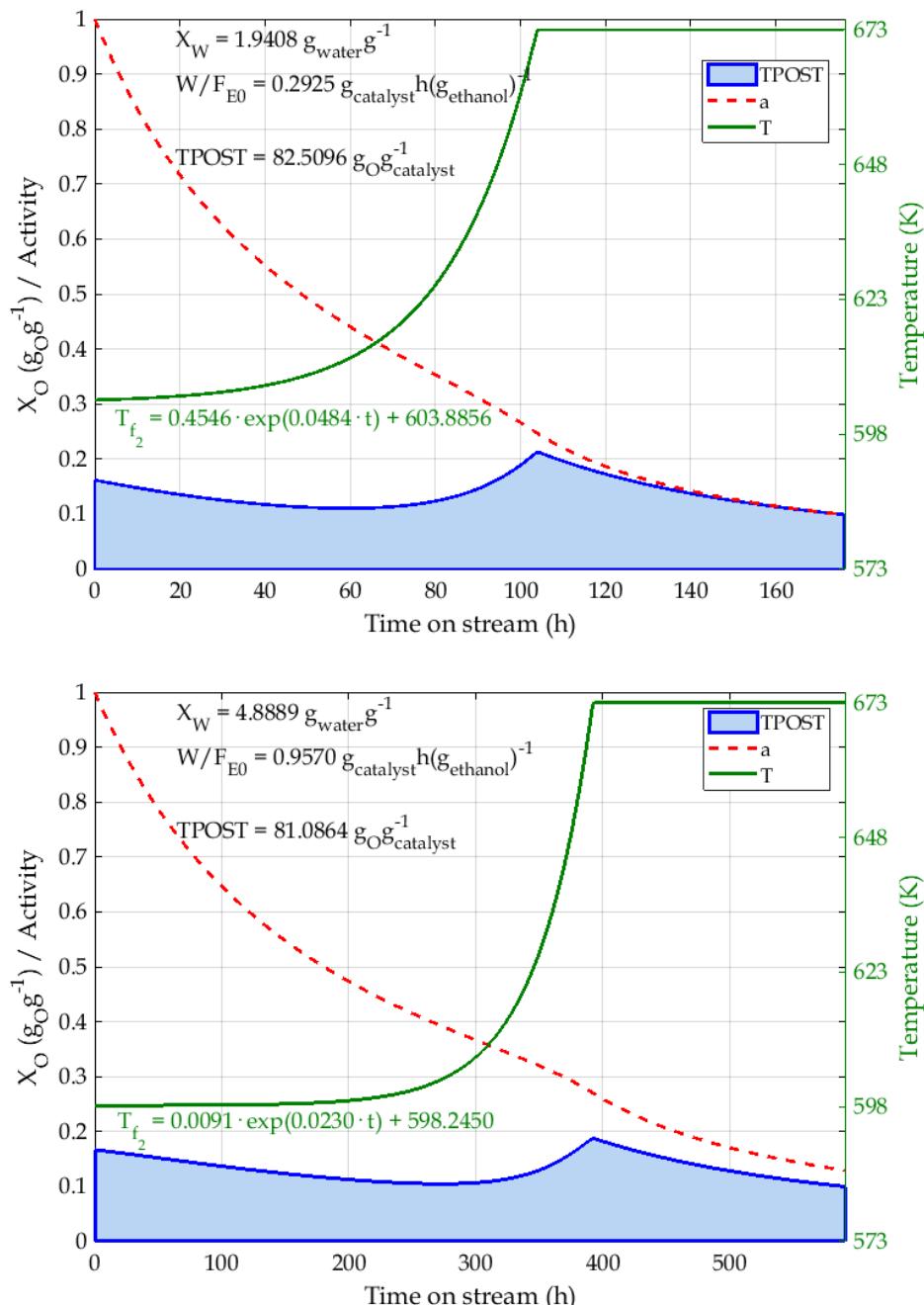


Figure B. 3. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_2 (Top: MECH solution; Bottom: HNN solution).

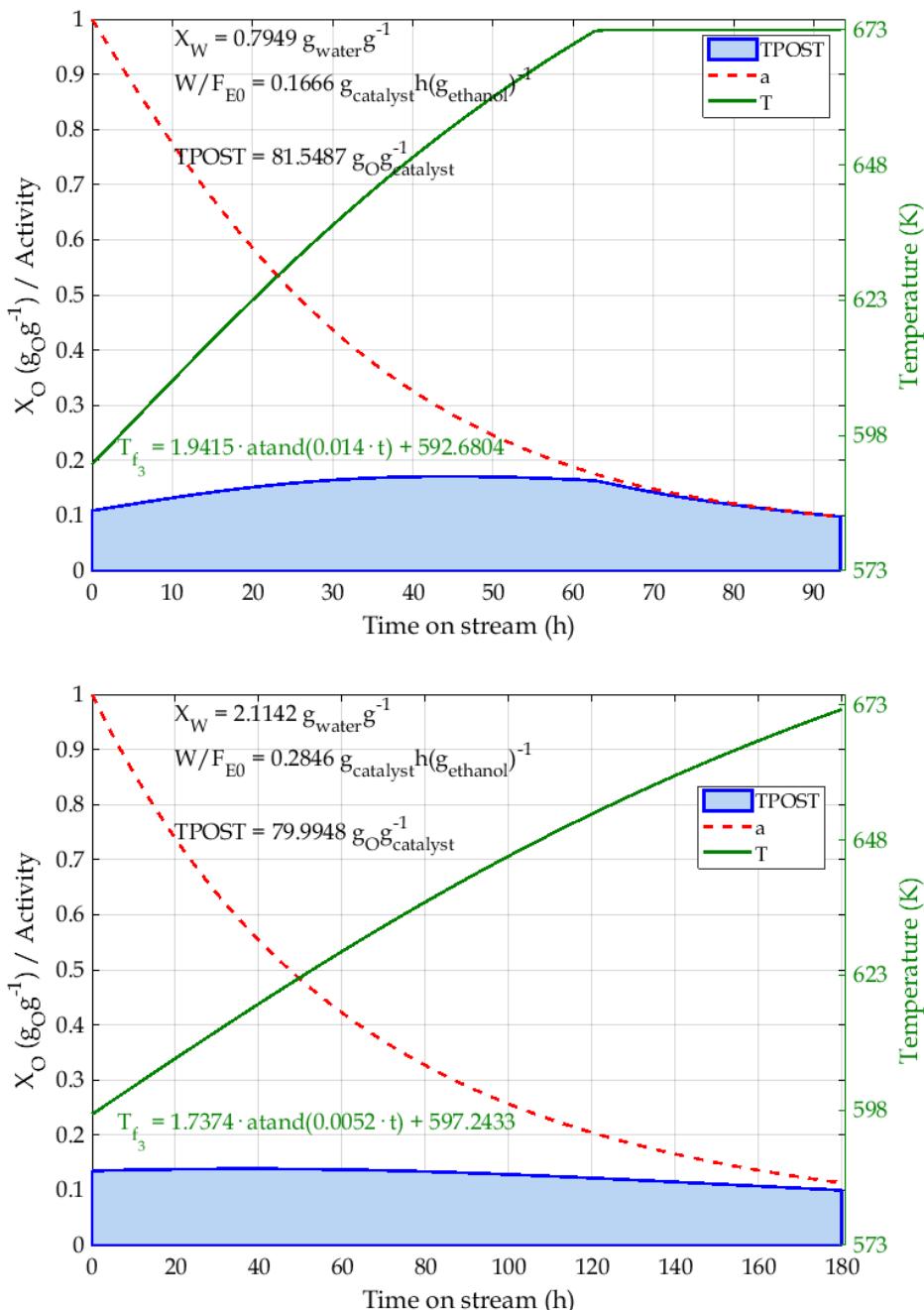


Figure B. 4. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_3 (Top: MECH solution; Bottom: HNN solution).

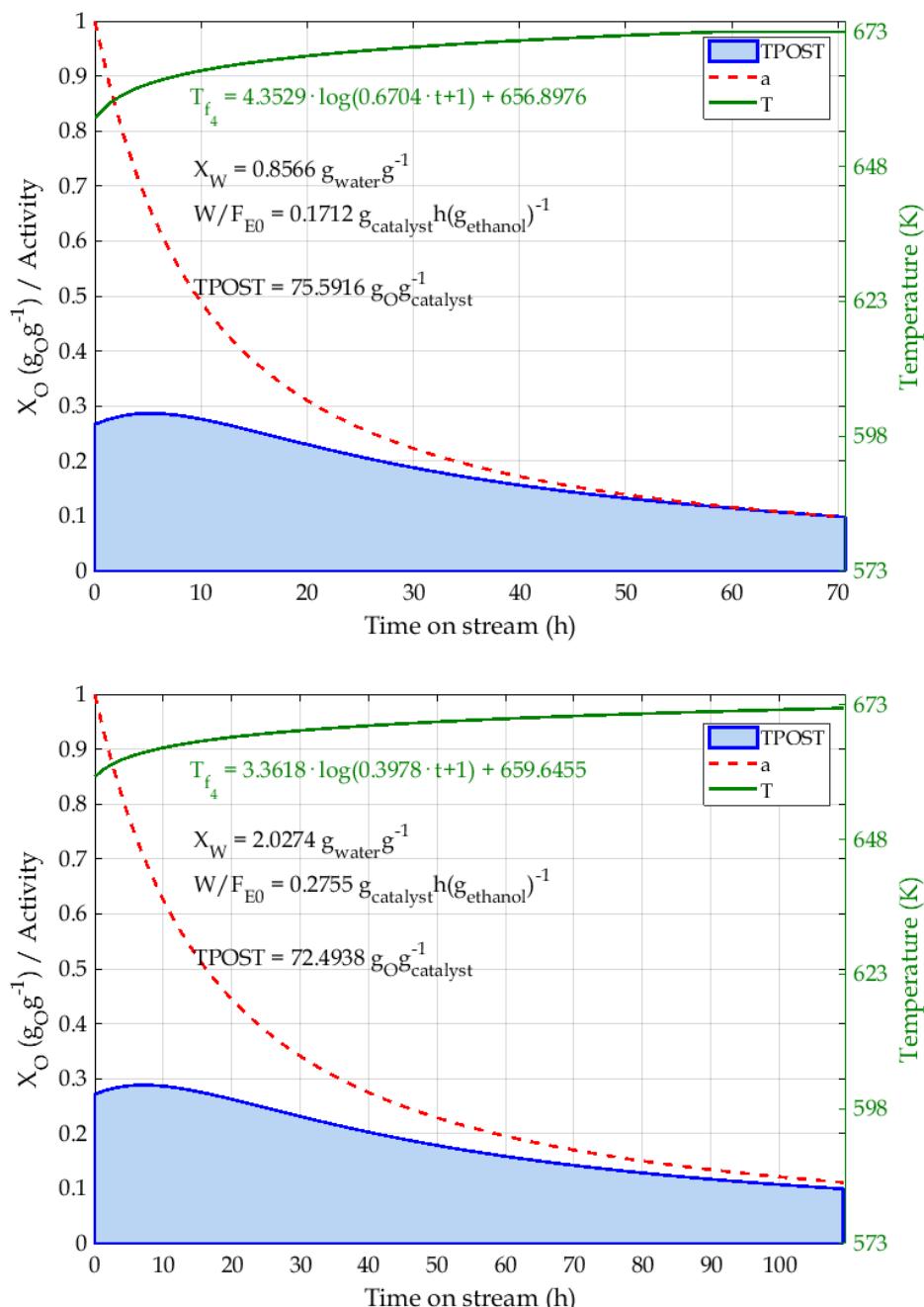


Figure B. 5. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_4 (Top: MECH solution; Bottom: HNN solution).

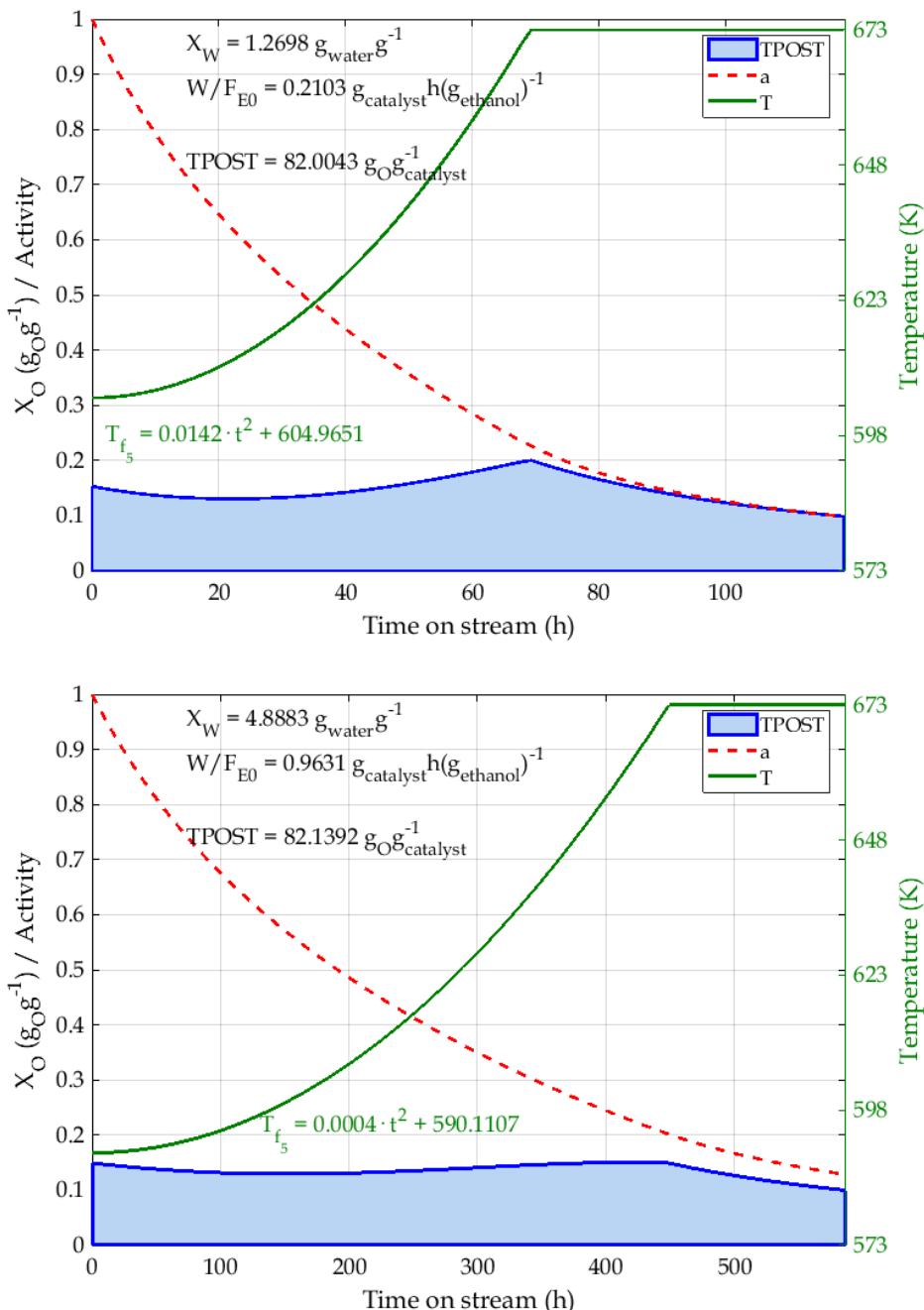


Figure B. 6. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_5 (Top: MECH solution; Bottom: HNN solution).

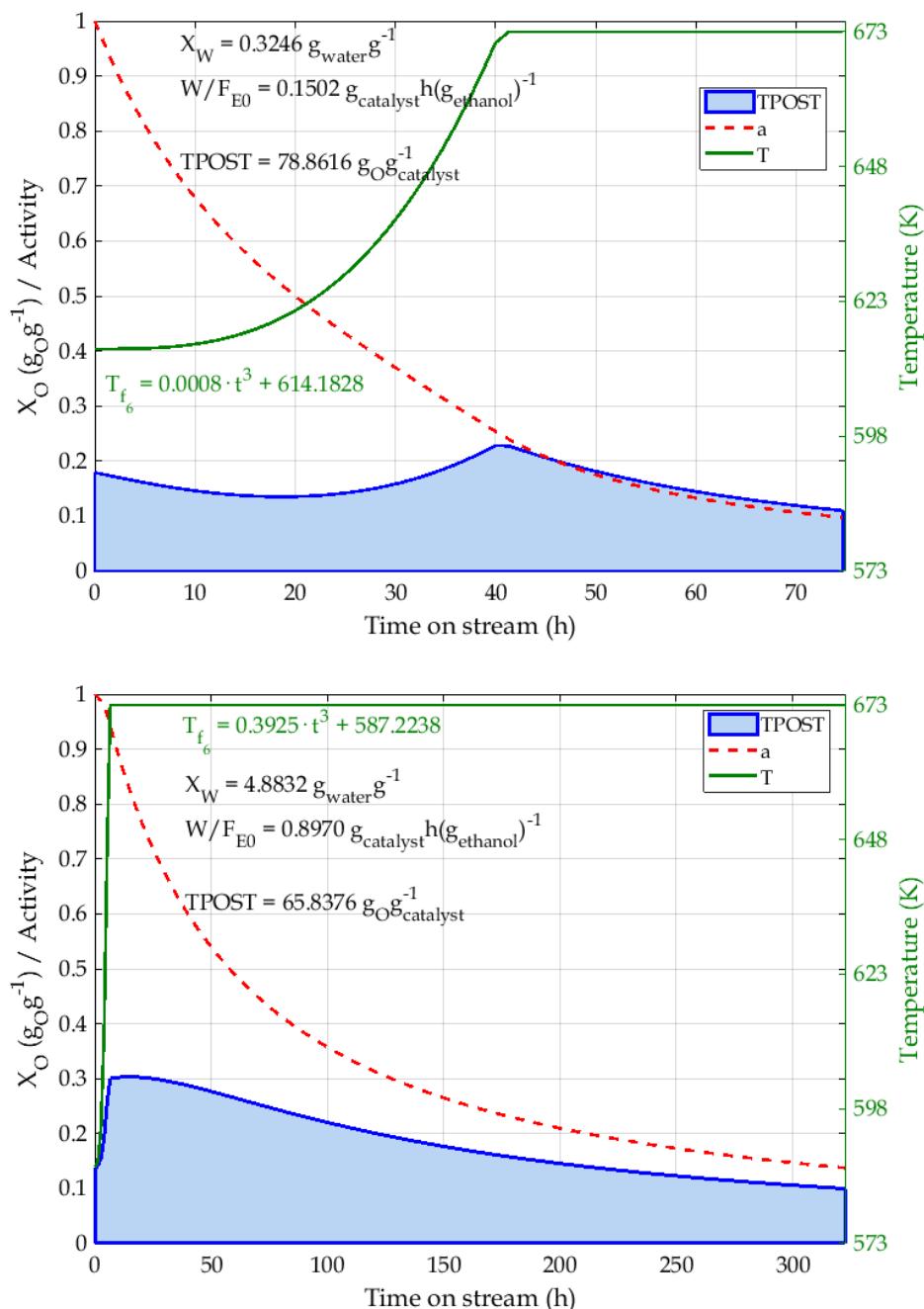


Figure B. 7. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_6 (Top: MECH solution; Bottom: HNN solution).

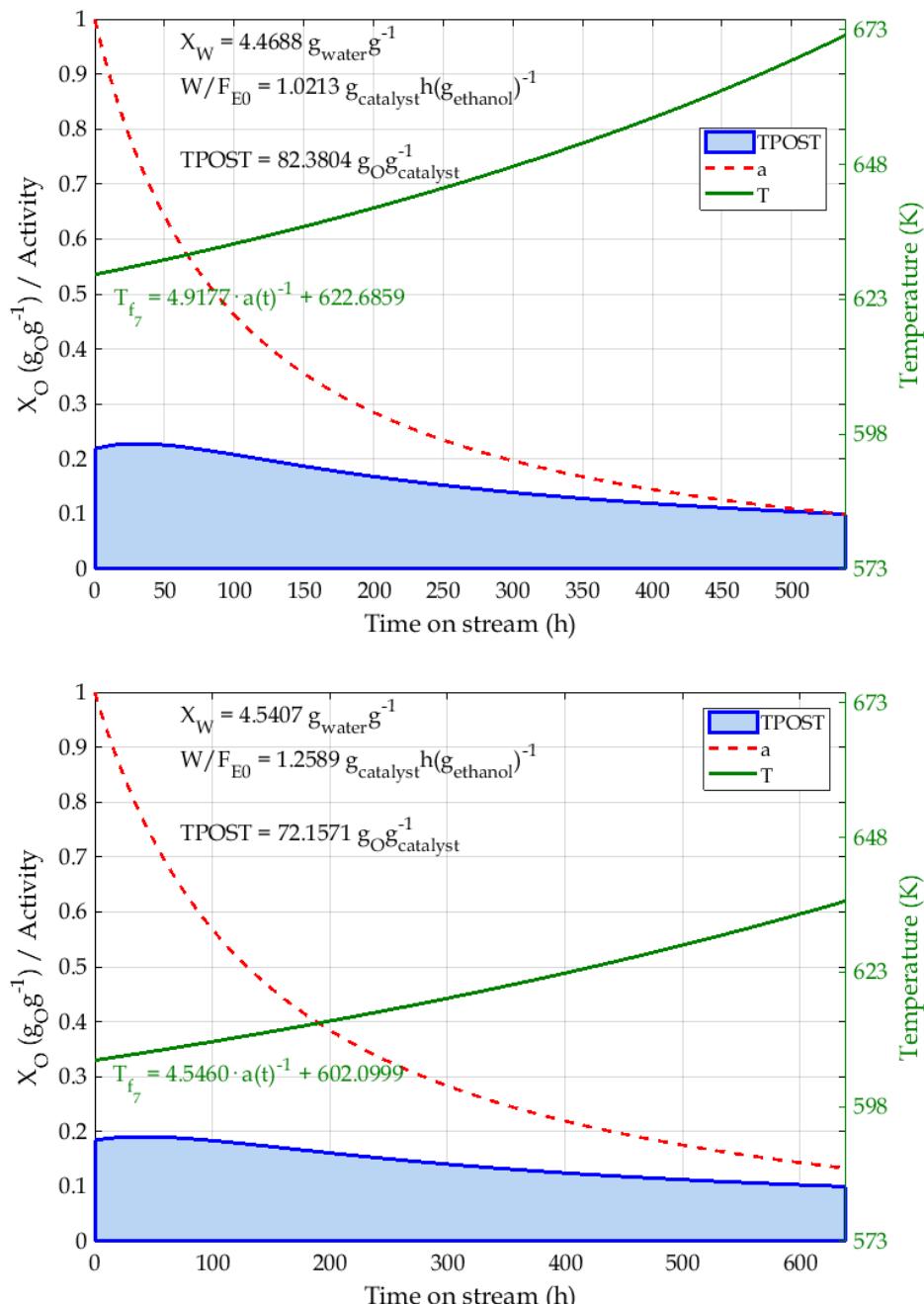


Figure B. 8. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_7 (Top: MECH solution; Bottom: HNN solution).

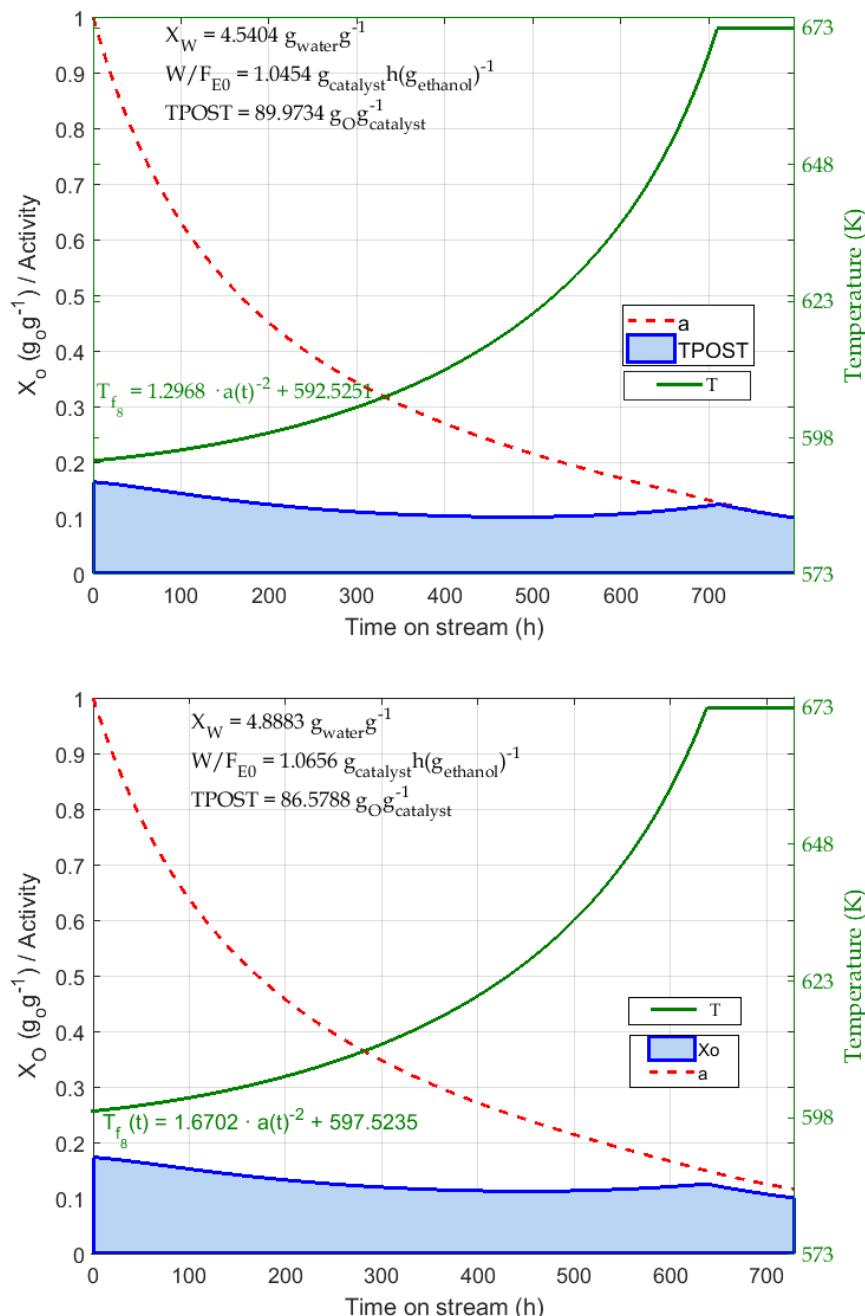


Figure B. 9. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_8 (Top: MECH solution; Bottom: HNN solution).

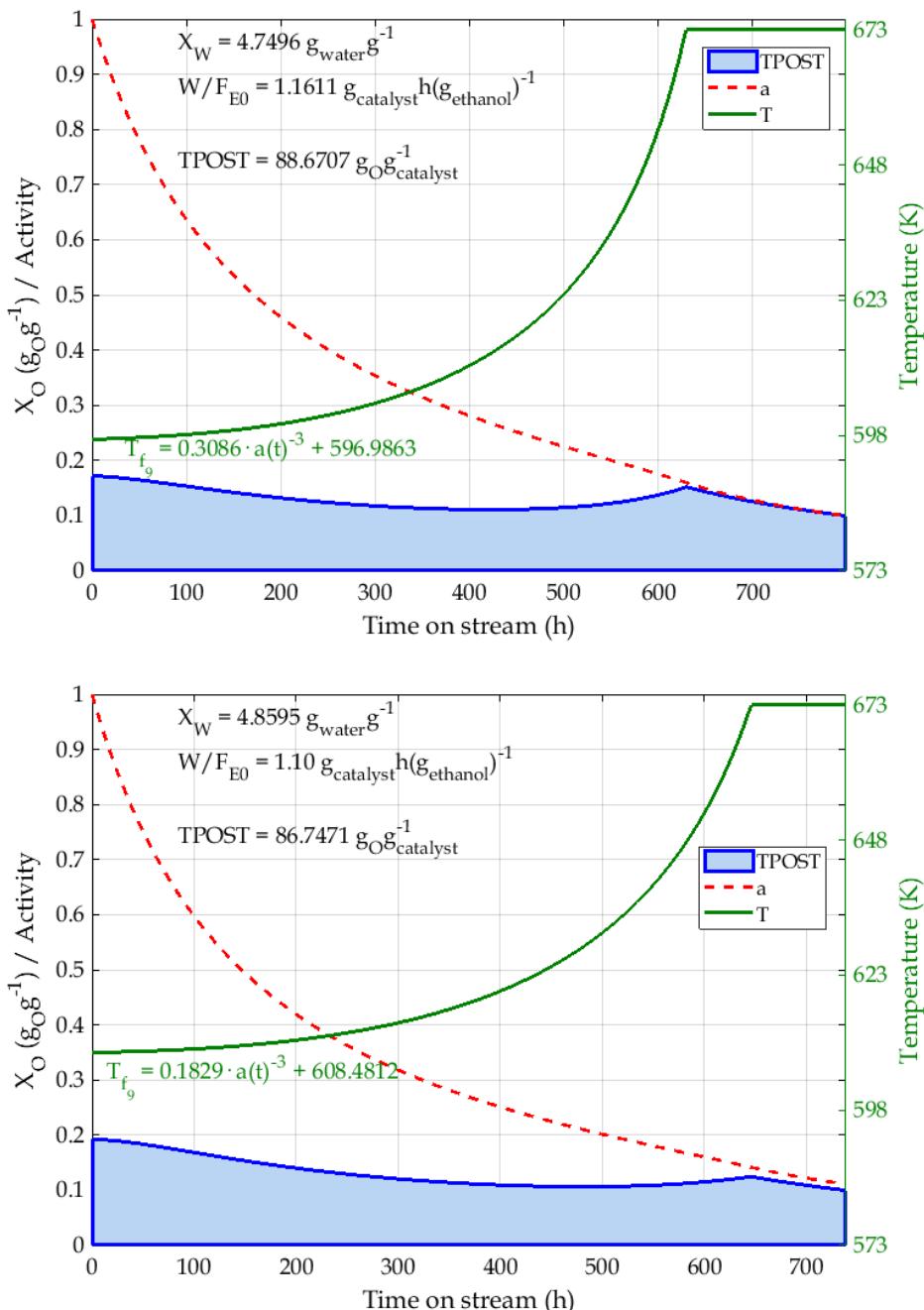


Figure B. 10. BTO process performance and operational conditions for the best solution obtained with the fixed-shape temperature trajectory f_9 (Top: MECH solution; Bottom: HNN solution).

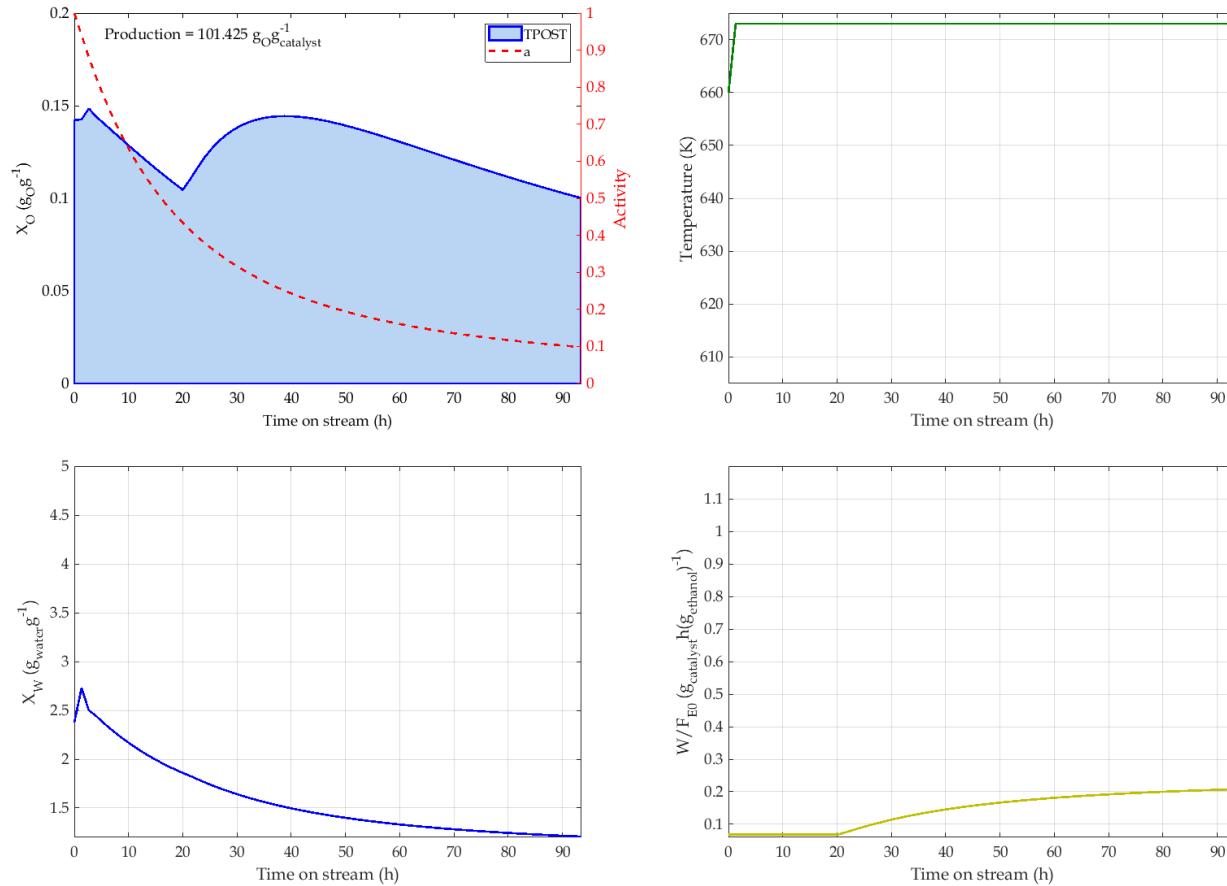


Figure B. 11. BTO process performance and operational conditions for the best solution obtained with evolutionary optimisation (n = 4).

Anex B: Optimisation Results

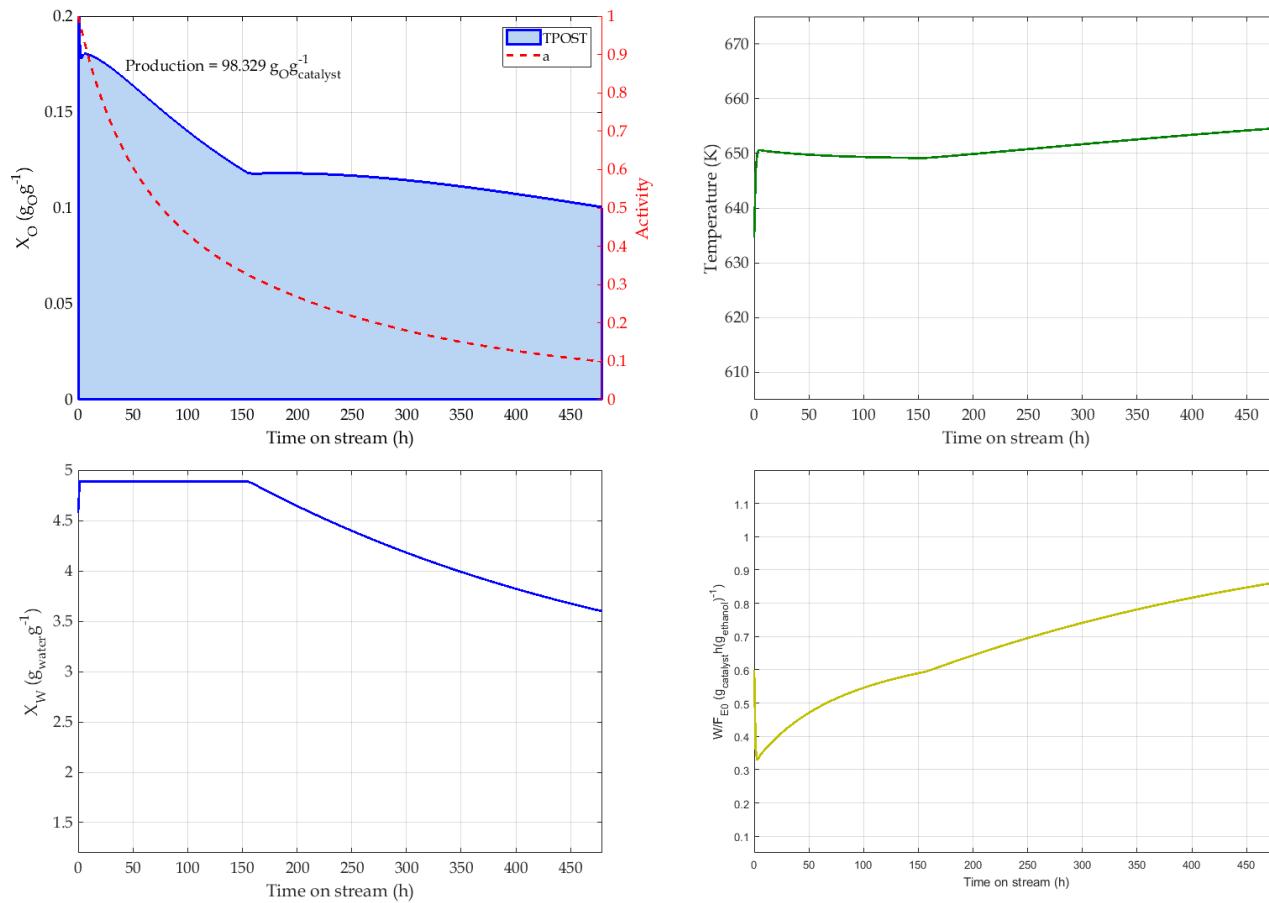


Figure B. 12. BTO process performance and operational conditions for the best solution obtained with evolutionary optimisation (n = 5).

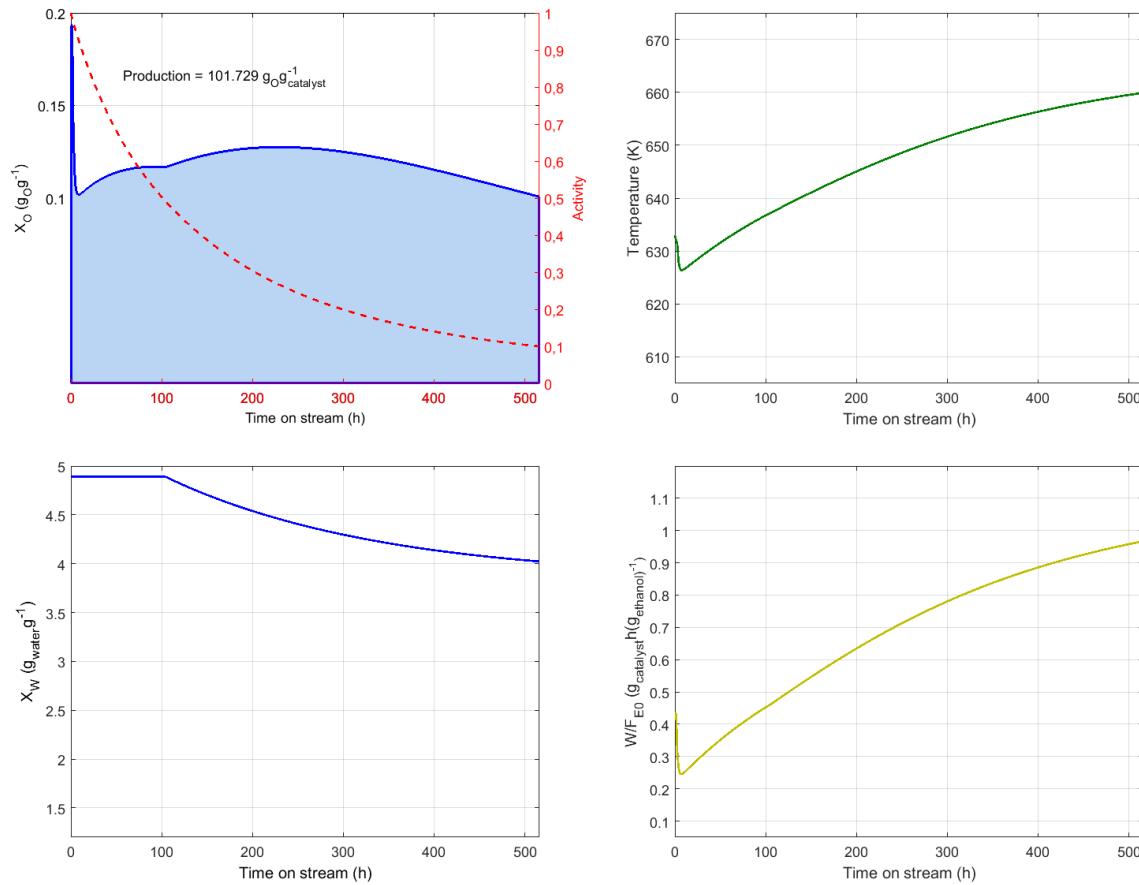


Figure B. 13. BTO process performance and operational conditions for the best solution obtained with evolutionary optimisation (n = 6).

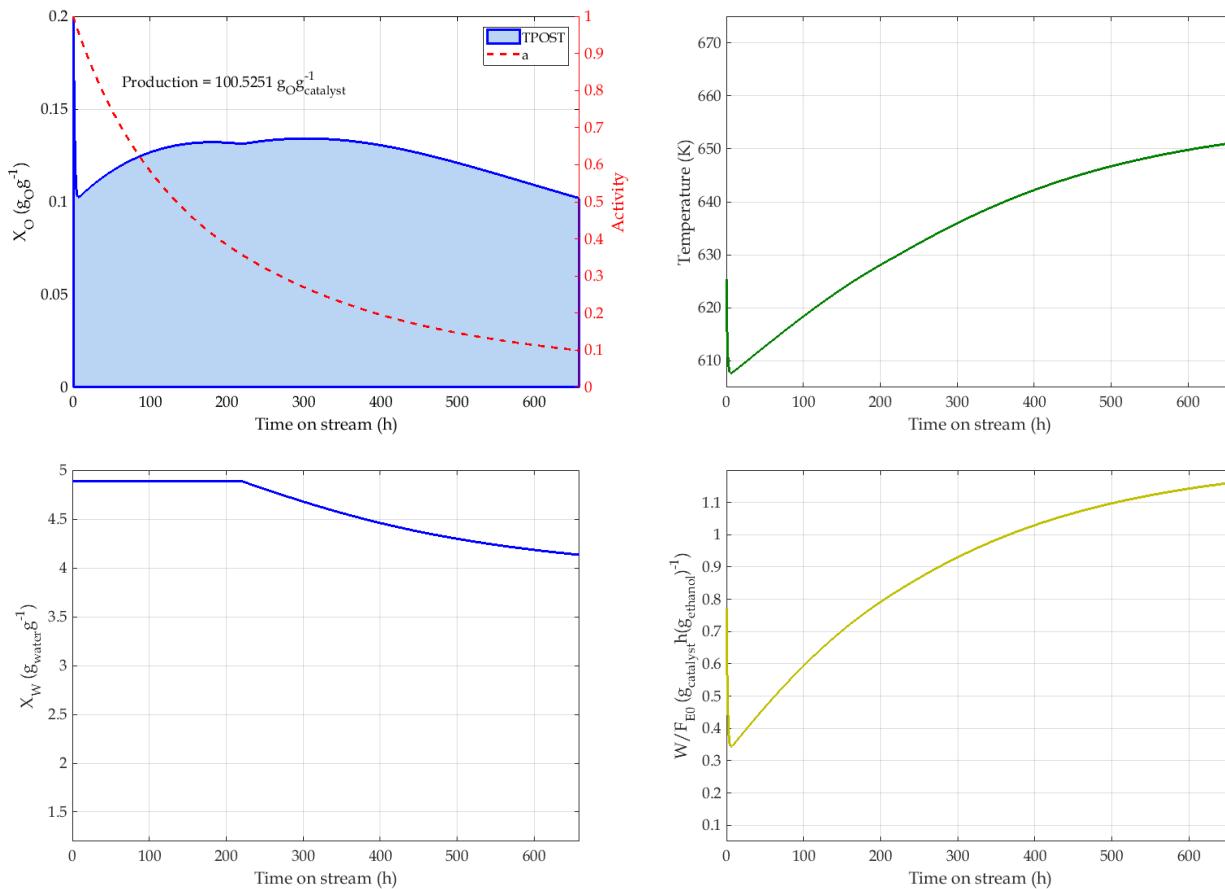


Figure B. 14. BTO process performance and operational conditions for the best solution obtained with evolutionary optimisation (n = 7).

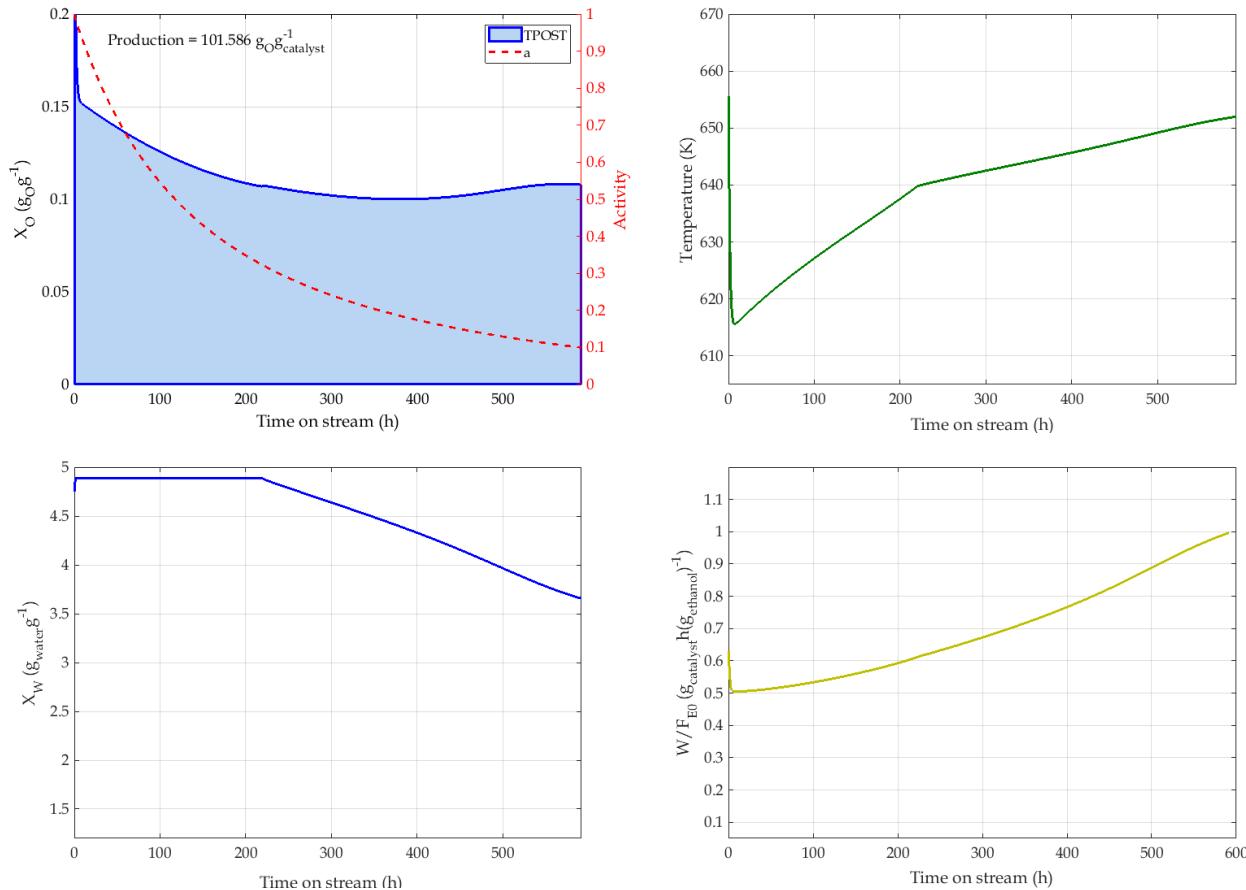


Figure B. 15. BTO process performance and operational conditions for the best solution obtained with evolutionary optimisation (n = 8).

