



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH

# **Use of Mathematical Methods in the Resolution of Chemical Engineering Problems**

A Thesis submitted as a compendium of publications by:

**Abel Valverde Salamanca**

In fulfillment of the requirements of the Doctor's degree

Supervised by:

Prof. Juan Jesús Pérez González

Prof. Francesc Recasens Baxarias

Department of Chemical Engineering

Terrassa, 2020

## Summary

1. Abstract .....	1
2. State of the art .....	3
3. Results and discussion.....	31
3.1. Binary interaction parameters from reacting mixture data. Supercritical biodiesel process with CO <sub>2</sub> as cosolvent. ....	33
3.2. Extraction of solid lanolin from raw wool with near-critical ethanol modified CO <sub>2</sub> —A mass transfer model .....	49
3.3. A Nonlinear Autoregressive Exogenous Neural Network Applied to Weibull Distributed Data. Lanolin Supercritical Extraction Model.....	63
3.4. Mathematical Modelling of Supercritical Fluid Extraction of Liquid Lanolin from Raw Wool. Solubility and Mass Transfer Rate Parameters .....	99
3.5. Assessment of the Conformational Profile of Bombesin by Computational Methods .....	127
4. Conclusions .....	139
5. References .....	143
Annex I.....	165
On critical properties from literature .....	165
On BIPs literature values and estimation .....	167
On Optimization Discussion.....	169
Annex II.....	171
Annex III .....	173
On Inlet Breakthrough Curve's Demonstration.....	173
On Weibull Distribution Results .....	180
On Solving Method Abstract.....	181

*To my family and friends*

## **Acknowledgements**

First, I would like to thank my department coworkers, especially Pilar Cortés and Margarita Murillo, for giving me the chance to carry on my work inside the university, which has been fundamental for my maintenance during the elaboration of this thesis.

I would like to also acknowledge the support of my coworkers and coauthors of the research papers included in this thesis, specifically Luigi, Patricia and Jesús. Without their advice and knowledge this thesis would not have been possible. Thank you Jesús for your continuous help and your friendship.

I am deeply indebted to my co-director Juan Jesús Pérez, who took me under his direction without hesitating, providing me with both his unconditional help in the most difficult last moments of this thesis, and his support in such an amazing field as protein folding is. My deepest thanks Juan Jesús for everything you have done for me.

Finally yet importantly, this thesis would have never even begin without the support of my tutor, co-director and mentor Francesc Recasens. Thank you Francesc, I debt you everything of these last wonderful years. Thank you for your trust in me, your daily support and for helping me to make this thesis real.

I want to dedicate this thesis to my family and friends, who have never stopped believing in me and have accompanied me through all the good and bad moments that have finally led me here. Thank you all.

## 1. Abstract

Historically, mathematical methods have played a key role in the development of natural sciences and specifically, in chemical engineering. This thesis consists of a compendium of five works that illustrate the utilization of selected mathematical methods to solve specific chemical engineering problems. Hence, the thesis is intended to cover both, a review of fundamental mathematical procedures for the solution of models raised from chemical phenomena, and a demonstration of their effectiveness to obtain useful significant results. However, the novelty of this thesis not only lays on the value of new parameters or conclusions previously unknown in the literature, but in some cases, the procedure applied to an specific problem, might be used as a reference for similar cases.

Optimization algorithms are frequently used in chemical engineering problems, since physical and chemical parameters from theoretical models are commonly fitted to experimental data. Here is where this thesis begins: a model of the supercritical methanol transesterification of triolein with CO<sub>2</sub> as cosolvent serves as a background to explore diverse local and global optimization algorithms. After obtaining novel binary interaction parameters (BIPs) for the Peng-Robinson EoS by means of optimization methods, the scope of the thesis moves to another recurrent topic in engineering problems: the resolution of non-linear partial differential equations (PDEs) systems. In this case, the issue is addressed from two different points of view: orthogonal collocation and finite differences methods. The former uses the study of lanolin extraction from raw wool with modified CO<sub>2</sub> under near-critical conditions as framework, while the latter explores the supercritical case. Although orthogonal collocation and finite differences are not groundbreaking methods, both the complex multiphase extraction model itself and the

resulting novel mass transfer coefficients can be seen as a significant achievements. Actually, the complexity raised by the presence of multiphase liquid lanolin under supercritical extraction is the basis of the following mathematical method explored in this thesis: neural networks. Thus, a hybrid model based on a non-linear autoregressive neural network and the Weibull statistical procedure is used to predict the extracted fraction values under diverse operating conditions. The potential of this method does not ends with the low predicting error obtained from such scarce available data, as it enables to determine the optimum extraction conditions too. Hence, the discussion of this method together with the previous three, provide the enough mathematical basis to approach the closing procedures explored in this thesis: Molecular Dynamics (MD) and Principal Component Analysis (PCA). These methods are aimed at assessing the conformational profile of bombesin peptide. The results suggest that the peptide adopts mainly helical structures at the C-terminus, with a significant role of the interaction between Trp<sup>8</sup> and His<sup>12</sup>, and less probable hairpin turn structures at the N-terminus. These achievements contribute to the determination of the most suitable binding conformation of the peptide, related to the interaction with its receptor that is involved in cancer development.

To sum up, this thesis approaches five different mathematical methods by means of five different practical chemical engineering applications, namely local and global optimization, orthogonal collocation and finite differences methods for PDE solving, neural network design for complex chemical modelling and, finally, Molecular Dynamics and Principal Component Analysis for peptide conformational profile assessment. In most cases, far from being a simple review of diverse mathematical methods applied to chemical engineering problems, the obtained results prove the novelty of both mathematical models and procedures developed in this thesis.

## 2. State of the art

Advancement in mathematics is closely connected to the development of natural sciences and, consequently, to engineering conforming an intricate symbiosis in which they feed to each other. This relationship has always been close and bidirectional. If we look at the development of natural sciences and engineering it is obvious that every step put forward in the development of any mathematical area has led to an upgrade of a specific science field or it has even been the starting shot for the maturation of a new physical area. This intimate relationship, where mathematics acts as a cornerstone necessary for further developments in natural sciences, comes from long ago.

This special relationship between mathematics and natural sciences might be under discussion these times, but in ancient times the origins of mathematics were born basically to explain the nature of the world. Thus, it is commonly accepted that about 20,000 years ago arithmetic and numbering were the first mathematical fields developed, obviously by the need of keeping track of transactions and events [1]. On the other hand, despite being developed after arithmetic, geometry also represented a significant advance for ancient cultures. Geometry was the result of a need of measuring lands, and contributed to build one of the most advanced civilizations of ancient times, with constructions that remain standing nowadays like the pyramids, temples and others. Although substantially developed by ancient Greeks, it was in ancient Egypt where arithmetic and geometry flourished [1]. However, it was in the times of the Pythagoreans, around 600 BC, when arithmetic and geometry that had for a long time been considered two completely different sciences that became part of the same umbrella: mathematics [1].

An important breakthrough in the history of mathematics came with the introduction of algebra by Al-Khwarizmi between 813 and 833 A.D., mainly after the publication of *The Compendious Book on Calculation by Completion and Balancing* [2]. We do not know whether Al-Khwarizmi took the concepts of algebra from Indian mathematics that like other ancient cultures such as the Chinese or the Babylonian had been using a primitive algebraic language previously, but his book was the first one to treat algebra as a mathematical discipline by itself. It is though necessary to give a special regard to *Arithmetica* (around 250 A.D.) by Diophantus that can be considered more a compendium of problems [3].

However, the introduction through Muslim Al-Andalus of Al-Khwarizmi contributions marked the main success of his work, achieving a huge impact in an Europe by that time still anchored in using geometry as the means of solving any mathematical problem. In fact, Algebra was not the only essential mathematical tool that Arabs introduced in Europe, since the number 0 arrived to Europe about the same time and through a similar way. The Arabs adopted the Indian numerical system, partly influenced by the work of Al-Khwarizmi [4] among others, and took it to Europe that was hindered in using the Roman numerical system [5]. This contribution along with the Algebra clearly supposed a highlight not only for the mathematical development, but also for natural sciences development for the next half of the middle age, as it is difficult to imagine any of the later technological developments after these days without algebra, algorithms or even our numerical system. These early contributions set the basis for mathematics developments that became crucial for natural science advancement.



Taking a leap forward into the XVII century, the development of the Analytical Geometry is an important milestone that deserves a comment. This was a great achievement that can be considered as the most important contribution to modern geometry since the publication of Euclid's *Elements* [6]. René Descartes in his *Géométrie*, published as an appendix to the *Discourse on the Method* (1637), dealt with a new method to characterize curves with algebraic expressions. This is one of the most important contributions by Descartes, besides others like the use of the Cartesian axis [7], a tool that can now be seen in most of graphics used in any field of natural sciences. However, it is unfair to refer to Descartes without mentioning Pierre de Fermat, who also did an outstanding work in this area being as relevant as the Descartes one was. Actually, it is well documented the correspondence between them, as well as of a historical controversy to be fair, which is not going to be described here [8]. Both studied the analytical solution of fourth degree equations, following the work of Gerolamo Cardano [7], who a few years earlier discovered a method to solve third order equations analytically. The so-called *Cardano's method* proved to be useful for solving for example, cubic Equations of State (EoS) of gases and supercritical fluids, improving the accuracy of the results obtained using numerical methods. Actually, this was an issue of scientific discussion [9] [10] in the past and is the method used to solve the Peng-Robinson EoS in the first work presented in this thesis [11]. In the first half of that work, it is stressed that it could have not been possible to obtain a proper solution without all mathematical contributions described so far. The work essentially aims to model the production of biodiesel from vegetable oil triglycerides reaction with methanol in presence of supercritical CO<sub>2</sub>. For this purpose, a complex reaction kinetics mechanism is proposed enabling to consider the thermodynamic state evolution including the generation and consumption of intermediate products. This last highlight is a new approach for a reaction mechanism already well

known in the literature as for example, in Osmieri et al. [12], Saka & Isayama [13], Campanelli et al. [14] and Maçaira et al. [15], among others. Taking into account the thermodynamic state of the species inside the reactor requires solving the cubic Peng-Robinson EoS by means of Cardano's method, for which Binary Interaction Parameters (BIPs) must be found, since for most of species those are unknown in current literature. The efficacy of the Peng-Robinson and other cubic EoS like van der Waals or Redlich-Kwong in supercritical mixtures has been shown in other works like those by Brunner et al. [16], Jaubert & Coniglio [17] and Reid [18]. Furthermore, both Brunner and Jaubert [19] [20] along with other co-workers developed software or predictive methods to find out BIPs for a wide range of gases and compounds, and even Gros et al. [21] developed a contribution method for this purpose. However, as far as we know, there has been only two attempts to find out the BIPs from data of a reacting supercritical mixture. On the one hand, Bertuccio's group [22] fitted BIPs related to an hydrogenation of a pharmaceutical organic intermediate to their results, and on the other, the group of Santacesaria et al. [23] who reported a new approach to find out BIPs from UNIFAC EoS.

The mathematical method for solving the system of ordinary differential equations (ODE) obtained as the result of the kinetic model described above requires to analyze again the relationship between natural science and mathematics and how decisive has been one to the other. In this case, the revolution came with the development of the differential calculus by Isaac Newton and Gottfried Leibniz in the XVII century. Avoiding the temptation to enter into the large discussed controversy of who developed these concepts first, what it seems clear is that both made remarkable contributions to it. Actually, it is probable that Newton came first with the idea of the differential calculus [24] with the work "*fluxions*" [25], describing the curvature of plane curves that he later

developed more extensively in the lemmas of his notorious *Philosophiæ Naturalis Principia Mathematica* [26]. However, the notation used by Newton and the procedure followed to make his demonstrations, based in classic geometry (powered by analytical geometry developed by Descartes and Fermat half century before) are much more difficult to read from the present perspective than those developed by Leibniz to whom we owe most of our present notation (with especial regard to Lagrange's notation too) and the algebraic procedure to operate in differential and integral calculus [27] [28]. Besides this specific issue, what is remarkable is that it might be the first of a few cases where mathematical developments were conceived as a consequence of the physics behind. In the words of R. Thom, taken from his chapter *The Role of Mathematics in Present-day Science* from the book *Logic, Methodology and Philosophy of Science VI* (1979) [29]: “*In most cases of great progress in physics, the body of mathematical tools anticipated the physicist's needs, and it very seldom happened that the mathematician had to create a new theory to satisfy the needs of the physicist*”. Indeed, Newton in his *Principia Mathematica* sat the basis with his lemmas for the development of mechanics, including the gravitational theory among others. This was the starting point for mathematicians to develop the differential calculus [29]. Newton's methods represent the basis of many mathematical procedures developed subsequently in a wide range of diverse fields. Some of them are discussed below, since they are involved in the development of the works included in the present thesis.

The first mathematical procedure developed after differential calculus found in this thesis is the Euler's method for solving ODE systems, developed by the mathematician Leonhard Euler. Euler was one of the most prolific mathematicians ever and, like Newton, responsible of another of those few cases where the need of solving a real problem results

in a new mathematical field. Euler was interested in solving the well-known “*The Seven Bridges of Königsberg Problem*” which set the ground for the birth of the Graph Theory [30]. Back to Euler’s method for solving ODEs, it is based on the definition of derivative and is the starting point for many ODE solving algorithms [31]. Specifically, in the first and second works included in this thesis there is a ODE system that is solved using MATLAB software, using the Dormand-Prince method algorithm from Runge-Kutta family [32], a 5th order ODE solver based on an early Euler’s method.

Taking a leap further in time, it was at the beginning of XVIII century that the time for solving an old controversy arrived. For years, mathematicians had been wondering whether the Fifth Postulate of Euclid -also known as the Parallels Postulate- was true or not. The answer to this question not only led to the origin of a new mathematical field, but I would not hesitate to say that it produced profound consequences in natural science and in our present daily life. History of mathematics is full of controversies and here there is another one between the Russian Nikolái Lobachevski, the Hungarian János Bolyai and one of the most significant mathematicians ever, the German Carl Friedrich Gauss. Besides who was the one who demonstrated first that the V Euclid’s postulate is false, it seems clear that Gauss first came up with the suspicion that it was wrong [33] [34] [35] [36]. However, Lobachevski and independently Bolyai first published in the late 1820s and in the early 1830s a work about a new geometry, where the sum of the angles of a triangle was less than  $180^\circ$ : the hyperbolic geometry was born [34] [36]. This development immediately raised the question: and, why not more than  $180^\circ$ ? This question was answered few years later by Bernhard Riemann to enlighten the elliptic geometry [34]. Nevertheless, the work of Riemann had a much broader scope. After his studies about the curvature of surfaces, taking the works by Gauss as a reference, he

demonstrated that Lobachevski's and Riemann's geometries could be represented by the same quadratic functions with a negative curvature in the former case and a positive one in the latter, being the Euclidean Geometry a limit case [37]. This geometry, known as Riemann's Geometry, and more significantly after the contributions of Levi-Civita to the curvature tensor and the work of Albert Einstein [37], supposed a major contribution to establish the General Theory of Relativity, concluding that the geometry of the space is Riemannian [34] [33]. Moreover, using the curvature tensor Einstein explained the geometry of space-time, (specifically, the Minkowski space, a non-Euclidean geometry [33]), being this a clear example that highlights how mathematics contributed significantly to the development of natural sciences and, concerned about the great impact of Einstein's most famous theory, to the way we live [38]. In the words of D. Dieks about this geometry: *"this branch of mathematics proved to be of decisive importance in achieving a break-through in his struggle for a relativistic theory of gravitation"* [39]. Another field where non-Euclidean geometry played an important role is in the field of optimization algorithms, regardless of being local or global. For instance, there are local optimization algorithms like the Gradient-Descent (GD) method or the Newton-Raphson algorithm that are based on Newton's method for solving equations, where the derivative is needed to converge to the lowest point [40]. Although traditional GD algorithms use Euclidean distances to converge, some variants of this algorithm (like ADAM, ADAGRAD or AMSGRAD) consider a non-Euclidean form of the objective function to carry out optimizations with higher performance [41]. Other mainstream local optimization algorithms are the Trust-Region algorithms. Those algorithms take the Newton method as starting point, using the Hessian matrix for the evaluation of the direction to where the minimizer takes the following steps, which is approximated [42] [43]. However, they are more related to non-Euclidean geometry than GD algorithms are,

as the overview is that the form of the objective function near the starting point, mostly hyperbolic or elliptic, is modeled with a quadric surface from where a set of trial points (at a certain step near the starting point) are evaluated and then the algorithm moves down towards the minimum [42]. Despite using Euclidean distances in most of the cases, there are several applications that use Riemannian distances [44] [45]. These optimization algorithms are widely used in diverse fields of natural science and specifically, in chemical engineering, since in chemical modelling unknown parameters are usually found by fitting the model to experimental results, like shown in Mentel et al. [46] or Almagrbi et al. [47]. Specifically, the latter work has been inspiring for the first paper included in this thesis [11], since an optimization algorithm is applied in order to find out the kinetic parameters of a supercritical biodiesel production model. Even more, the objective function presented by Almagrbi et al. [47] has been a reference for the work developed by Valverde et al. [11]. Specifically, in the first work presented in this thesis there is a global optimization where both, kinetic parameters and Peng-Robinson's Binary Interaction Parameters (BIPs) must be found simultaneously since they are dependent one of each other. This leads to an iterative optimization algorithm with a significant number of parameters to fit (6 kinetic parameters and 14 BIPs). Therefore, the objective function form may be seen as an "egg-cup", namely, a non-Euclidean surface with several local minima but only a global one [48]. In the first paper [11], two potential solvers from the MATLAB software were considered: the *GlobalSearch* and the *Genetic Algorithm* [49]. The former applies a chosen local optimizer, like those already explained above, at a different starting points distributed throughout the possible solutions domain (this is a simplification of the *basins* concept used for the election of the initial points [49]). Despite the stiffness of the problem, the former demonstrated the best performance. Genetic Algorithm is, again, one of those tools where natural science and specifically,

genetics, leads to a new mathematical field. Genetic Algorithm (GA) is based on the evolution of biological systems first proposed by Charles R. Darwin in his masterpiece *On the Origin of Species* published in 1859 [50]. This is a great example how science led to the *evolution* of mathematics, inspiring John Henry Holland in 1975 [51] to develop a numerical method where an initial population is randomly chosen and structured in strings: the “chromosomes”. Those are then evaluated with the objective function, the better performance the higher probability to be selected for an intermediate generation. Finally, a crossover and mutation of the data contained in the strings of this intermediate generation arise a new population of “chromosomes”, more likely to accomplish the objective function, starting a new iteration until convergence [52]. It is a stochastic algorithm with high robustness and non-derivative, unlike the other local algorithms here discussed and used by GlobalSearch, which means that not continuity nor differentiability is required. Despite the higher computational cost and the lack of convergence guarantee, it is suitable for non-smooth problems [52] [53]. Both methods were considered in the first work presented in this thesis, and despite the stiffness of the problem, GA did not show a better performance, while it took a much higher computational cost. Consequently, *GlobalSearch* was selected as the best choice for solving the twenty-one parameter optimization [11].

Before going back to other contributions produced after the development of the non-Euclidean geometry, it is worth to discuss briefly about an emerging field related to optimization: the Artificial Neural Networks (ANN). ANNs might be one of most paradigmatic examples connected with the scope of this thesis: the *quid pro quo* relation between natural science and mathematics. In 1943, the neurophysiologist and psychiatrist Warren McCulloch joined Walter Pitts, who had studied logics and biophysics and

published *A logical calculus of the ideas immanent in nervous activity* [54], a work where a first mathematical model of neuron activity was proposed. It was strongly influenced by the famous Alan Turing's machine [55], early theoretically developed in 1936 and highly improved by the cryptographic efforts to break German Enigma's machine [56]. The birth of this field was responsibility of another scientist, the psychologist Donald Hebb [57], who in 1949 contributed with the first notions of Unsupervised Learning [58]. McCulloch and Pitts neuron model might be seen as a discrete time dynamical system, where the synapsis of a neuron are activated depending on a specific synaptic resistance, namely a weight multiplied to input value and a transfer function modelling the Heaviside function that will fire output if weighted input gets over a certain threshold. Usual transfer functions are the logistic function or the hyperbolic tangent, among others [59]. An aftermath of ANN made the neural net evolve and improve (Perceptron, ADALINE, Multiple-Layer feedforward networks, Recurrent Networks, etc.), leading to a common architecture consisting of an input layer with one or more input neurons; none, one or more hidden layers; and a final output layer with one or more neurons. In both neurons, from hidden or output layers, all inputs are multiplied by a weight and mixed in each neuron with its own transfer function and threshold [60]. The goal of the ANN is to adjust those "synaptic" weights to obtain the correct output from a certain inputs, this is the learning step, thus once it is over the ANN will be ready to predict unknown values. The need of minimize the output error with the weights as parameters led to understand this process as an optimization, including then a backpropagation routine, where weights were optimized with a local optimizer like the Gradient Descent (GD) method [58]. Improvements in this routine were critical for the performance of the ANN, leading to advanced GD methods, being one of the most used in present days the Levenberg-Marquard (LM) algorithm. This algorithm can be seen as a precursor of Trust-Region



(TR) algorithms indeed, as its mainly achievement, among others, is adding a parameter to ensure the convergence of GD, the same function Hessian matrix performs in Trust-Region algorithms [43] [44]. This upgrade, in addition to later improvements like its recursive version, make the LM algorithm one of the most efficient algorithms for a better convergence in ANN [61]. This was the algorithm chosen for the training phase of the ANN constructed and described in the third work presented in this thesis [62]. The goal of this third work was to model the supercritical extraction of liquid lanolin from raw wool. However, since the liquid phase is a mix of two different solubility lanolin fractions, the possible mass diffusion phenomena, not only between supercritical and liquid phases, but inside liquid phase fractions too, become difficult to model [62]. Previous works on supercritical extraction of lanolin from raw wool are available, like those from Cygnarowicz-Provost et al. [63], Kuo et al. [64] and Jones et al. [65]. However, none of them present the mixed fraction reported by Eychenne et al [66]. In this kind of situations, ANN can be useful, since they make possible the prediction of extraction results at a different input operation variables without the need of knowing the mechanisms that occur inside the extraction column [67]. For this purpose, two major references were considered: first, the work with adsorption columns by Bulsari and Palosaari [68] [69], who designed a feed-forward neural network using Levenberg-Marquard learning algorithm to predict the breakthrough curve at the columns exit. Second, the work by Fullana et al. [67], who designed a feedforward multilayer neural network trained with backpropagation algorithm to predict the breakthrough curves of black cumin seed oil supercritical CO<sub>2</sub> extraction. In the case reported in the third paper of this thesis, the ANN have been applied to a time serial of lanolin extracted values, previously regressed with a Weibull probabilistic distribution. This distribution, named after Waloddi Weibull and developed in 1939 (despite Fréchet identified the distribution

in 1927 was Weibull who first gave its meaning and utility), has several applications such as material reliability, geophysics or social sciences, among others [70]. Therefore, a non-linear autoregressive exogenous neural network (NARX) was applied, where previous outputs of the ANN are used as input for the following time prediction, along with other external (exogenous) inputs [71]. Those ANN were first introduced by Lin et al. [72] in 1996 after the Time-Delay Neural Networks were first published by Lang et al. [73] in 1990, and have demonstrated to be very useful in fields like engineering [74] or economics [71], mostly for forecasting unknown future data from previous known time series.

These new methods would have not been possible without a solid background in mathematical tools, like those developed in the previous centuries. Nevertheless, those developments exposed so far only laid the bases to new ones that had to come in revolutionary times, in the early XIX century. Newton and Leibniz's differential calculus made possible a more accurate model of physical phenomena, so once again science had an assignment for mathematics: to develop the tools for solving those physical models. French Joseph Fourier, in order to solve the heat flow equation as well as the vibration equation, developed a method to transform the corresponding differential equations into a series of periodic functions, not only reaching a correct solution, but contributing with a new integral transformation: the Fourier Transform (FT) [75] [76]. R. Thom agrees in *The Role of Mathematics in Present-Day Science* [29] that the development of FT was completely motivated by the need of solving those physical equations, so mathematical development was purely a natural science achievement. But benefit was not only for heat flow and vibrations, since Fourier's analysis was a milestone for Optical Signal Processing, field upgraded after Fourier solution applying the FT to D'Alembert's wave

equation [75] [77]. However, Fourier Transform is best suited for solving differential equations on domains ranging  $(-\infty, \infty)$  where no boundary values are taken into account, reason why is commonly applied to diffusion problems in infinite volumes. Furthermore, the FT solution is a periodic function, which might be an inconvenience depending on the situation [78]. For solving differential equations in a positive time-domain obtaining non-periodic solutions, a different integral transform was developed by a Fourier coetaneous: Laplace, with the Laplace Transform (LT) [78]. Actually, Pierre-Simon Laplace together with Lagrange and Legendre judged Fourier's work, receiving a very critical assessment by the lack of rigor in their opinion [75]. Although the LT was not initially developed to solve any particular physical phenomenon, it has been used in several fields. For instance, for solving one of most important equations in the history of science: the time-dependent Schrödinger equation (although the most common way to solve it is considering Schrödinger equation as an eigenvalue problem) [79] [80]. Among other notorious contributions, the theory of probability developed by Laplace sat the foundations of modern probability. As Fourier said: *"Laplace has gathered and established the principles. So, it became a new science, subject to a single analytical method, and of a prodigious extent. Fertilizes in usual applications, it will one day shine brightly light over all branches of natural philosophy"* [81]. The influence of Laplace's work can be easily identified in the fourth work presented in this thesis inspired in a previous work on ANN modeling [62] that involves a physical model of the supercritical extraction of liquid lanolin from raw wool is presented [82]. The main impediment to develop a physical model in the previous work was the complexity of the liquid lanolin phase that consist of a diverse solubility lanolin mixture. In the fourth work this is solved by a simple hypotheses inspired by the work of Puiggené et al. [83], a main reference for their analytical solution of a two coupled partial differential equations (PDE) system, using

Laplacian domain and incomplete gamma function resolution. By this means, Laplace Transform is applied to one partial differential equation, leading to an integral equation that is numerically solved. Thus, the PDE system is solved by the finite differences algorithm (see the impact of Newton and Leibniz contributions, another algorithm based on derivative definition [84]). Finally, a global optimization with genetic algorithm is carried out in order to find out the mass transfer coefficient and the equilibrium constant [82].

More recently, between the end of XIX century and early XX century a new singular relation between mathematics and natural science appeared to change the conception of the world: functional analysis and quantum theory. The beginnings of functional analysis relays on the axiomatization and generalization of Euclidean geometry performed by David Hilbert, published in his 1899 work *Grundlagen der Geometrie* [85]. After the contributions of Gauss, Lobachevski, Bolyai, and Riemann to the non-Euclidean geometry, Hilbert looked for a formalization of the geometry to break with Euclid's *postulates* and generalize these diverse spaces [86]. Thus, Hilbert presented a generalization of Euclidian geometry, a complete metric space whose norm is given by the Euclidean scalar product [87]. Hilbert space generalizes the concept of *vector* to objects other than "*arrows*", as for example functions or series as well, and extends Euclidean traditionally finite dimensions to infinite space dimensions [88]. This treatment of vector spaces defined by a norm is the essence of functional analysis. However, what really made the difference was its axiomatization of geometry, as Hilbert space is structured in twenty-one axioms that not only sat the basis for later geometrical treatments, but also induced the same axiomatic method to physical sciences. The relation between Hilbert spaces and physics was conceived at the University of Gottingen, home

of great mathematicians like Gauss, Riemann and Klein, and where, by that time, Born, Heisenberg, and Jordan among others, carried out their research in quantum mechanics [86] [89]. Hence, it is clear that the development of functional analysis was very interrelated with the physics of the early XX century, leading to the *23 Problems* proposed by Hilbert at the Second International Congress of Mathematics in 1900, mostly of mathematical nature except for the sixth one: *Mathematical Treatment of the Axioms of Physics* [86] [89]. This assignment was not actually performed by Hilbert, since John von Neumann is regarded as the developer of the axiomatization of quantum mechanics and even of the same Hilbert space notion [89] [90].  $L^p$  spaces were developed by Lebesgue and, later, Banach completed the normed space, generalizing  $L^2$  Hilbert space (Banach space norm must not be an inner product as it is in a Hilbert space) [87] [91]. The generalization proposed in Banach's contribution and its capability to include numerous sub-spaces, made the Banach space a central tool in functional analysis with numerous applications still today [92]. Actually, the study of quantum mechanics operators, like the Hamiltonian operator in Schrödinger equation, led von Neumann to develop in 1930s an algebra of operators that was indeed a Banach sub-space. This algebra, along with the  $C^*$ -algebras, are called *operator algebras*, and have been developed mainly by mandate of quantum mechanics, in order to find the proper tools and spaces to describe and solve the generated expressions [89]. A central theorem for the development of von Neumann's algebra was his spectral theorem for the definition of linear operators, which was further development in the Hilbert's spectral theory [89] [93]. However, the fundamental concept of eigenvalues and eigenfunctions in spectral theory were originated years before Hilbert built his theory (despite it was Hilbert who first used the terms eigenvalue and eigenfunction in the formalization of spectral theory [94]). Actually, it is generally attributed to Jacques Charles François Sturm and Joseph Liouville the origin of

eigenvalues and eigenfunctions, a crucial idea for spectral theory and functional analysis, for solving second-order differential equations [89] [95]. The first publication about the Sturm-Liouville problem dates from 1837, whether its real transcendence became after the Hilbert space was born, when in 1910 Weyl developed the Sturm-Liouville problem related to Hilbert space operators, eigenvalues and eigenfunctions and inner product of functions in  $L^2$  spaces [95]. The Sturm-Liouville problem describes the structure of a second-order differential equation that is commonly obtained in many physics fields. As far as the solution is concerned, Sturm-Liouville focus the equation as an eigenvalue problem, namely an hermitic operator applied to a function equal to the product between this same function and a constant. It results in a series of polynomials that are orthogonal in the Hilbert space, what means that the inner product between two of those solution polynomials is null for different polynomials and different to zero otherwise (the solution of this last case would be one if they are orthonormal) [96]. Hence, the Sturm-Liouville problem can be applied to diverse second-order differential equations proposed by previous mathematicians in order to model specific physical phenomena, raising in a different orthogonal polynomial series depending on the problem, each one represented by a power series expression or a differential equation expression (called the Rodrigues' formula), and bounded or unbounded in an interval domain [96]. Some examples include Hermite, Laguerre, Chebychev or Legendre polynomials among others. Those orthogonal polynomials together with Sturm-Liouville problem had a great impact in the quantum mechanics field, as for instance, the solution that Schrödinger gave for his fundamental equation for the hydrogen atom in his 1926 work *Quantisierung als Eigenwertproblem* focused in its identification as an eigenvalue problem [97]. After separation of variables, the angular equation of Schrödinger is, indeed, a Sturm-Liouville equation, more precisely a Legendre equation, the solutions of which are the associated Legendre

polynomials. The expression of these polynomials depend on two integer parameters  $l$  and  $m$  leading to the spherical harmonics. Schrödinger carried out a physical interpretation of these numbers and named them as the angular momentum number and the magnetic number, respectively [95] [98] [99]. Is it not amazing that an abstract mathematical theory developed by Sturm-Liouville could be used to make sense in the real world almost a hundred years later? Another application to physical science of orthogonal polynomials came in 1965, when Villadsen and Stewart published their work *Solution of Boundary-Value Problems by Orthogonal Collocation* [100], where they explain a new method that approximates the solution of a partial differential equation to a linear combination of orthogonal polynomials. Partial derivatives are transformed into total derivatives as orthogonal polynomials are collocated at different domain points, letting the polynomials coefficients only single-variable-dependent. The roots of the orthogonal polynomial with higher degree are usually chosen as collocation points, mostly for its better performance and simpler solution [100] [101]. The larger the series of polynomials is (the higher degree of the last polynomial), the larger number of collocation points and hence, the less error in the solution and the higher computational cost. In any case, the method exhibits a good performance even with few collocation points. This idea was used in a work by Bruce Finlayson [102], a key paper for the second work presented in this thesis, since Finlayson applies orthogonal collocation to solve an evolutionary partial differential equation system, raised from the modelling of a packed bed reactor. Finlayson work along with others, like Morbidelli et al. [103], where key references to develop the second work included in the present thesis by Valverde et al. [104]. In that work, the near-critical extraction of lanolin from raw wool is modelled, giving rise to two coupled partial differential equations depending both on time and extractor height, that are solved by orthogonal collocation. Legendre polynomials were

applied after redefining a set of non-dimensional variables in order to meet the orthogonal polynomials domain. The main difference in this modelling respect the third and fourth papers already discussed is that when CO<sub>2</sub> is at near-critical conditions, lanolin is a non-porous solid, so a physical model is presented considering two lanolin fractions with different solubility, as proved by Eychenne et al. [66] with no diffusional phenomenon inside the solid phase [104]. The developed model is based on the shrinking-sore concept for solid extraction presented by the Octave Levenspiel in his book *Chemical Reaction Engineering* from 1999 [105], where both the change on lanolin fraction and the lanolin layer end are modeled by a logistic function using the work by Fullana et al. [67] as reference.

A closely related field for the understanding of atomic world is Statistical Mechanics. This area has its beginnings even before Hilbert first contributions to mathematical and physical axiomatization, in the early second half of XIX century, when Ludwig Boltzmann and James Clark Maxwell first proposed a new approach to understand macroscopic magnitudes from statistically distributed microscopic states [106] [107]. Actually, whether foundations of statistical mechanics were further developed in 1902 in the work: “*Statistic Mechanics*” by Josiah Willard Gibbs [107] and later, in 1911 in the work “*Begriffliche Grundlagen der statistischen Auffassung in der Mechanik*” by P. Ehrenfets and T. Ehrenfest [108], Maxwell previously had published “*Illustrations of the Dynamical Theory of Gases*” [109] in 1860. In his work, Maxwell gave the first statistical interpretation of a thermodynamic system based on the distribution of velocities of its molecules. Boltzmann deepened more into Maxwell statistical approach coming up with the concept of “*ensemble of systems*”. Namely, each macroscopic system might be seen as an *ensemble* of different systems statistically



distributed [110] [107]. This idea approaches the macroscopic properties of a system from the point of view of the contribution of its atoms or molecules. Despite finding a harsh criticism by 1868 and 1871 when Boltzmann first published his results, it was decisive for XX century physics, in the “*quantum*” concept by Plank or the Einstein statistical approach to physics for instance [107]. In 1868, taking into account the Liouville Theorem (a key theorem concerning to the equation of motion) [110] [111], Boltzmann raised an important dilemma: the Ergodic Problem. It derives from the Ergodic Hypothesis, which can be described according to Cárcamo as: “*Given a dynamic system, we can associate several physical magnitudes. The average of the values that one of these physical magnitudes takes in only one of these systems, as it evolves over time, is equal to the value that we would obtain at averaging the same magnitude measured in many similar systems at the same time*” [112]. According to Gallavotti: “*The fraction of time in which the energy is in the shell  $dk$  equals the average number of particles with energy in  $dk$ : a forerunner of the ergodic hypothesis*” [110]. Namely, the ergodic hypothesis considers that the distribution of the different ensembles that a system presents over time is equal to the distribution of the different microstates present on a single given time. Rédei [90] consider this hypothesis and its aftermath as a clear example of mathematical and physical merging, as the attempts to demonstrate this hypothesis, clearly derived from the modeling of a physical phenomenon, led to the development of the ergodic theory. After Boltzmann hypothesis, ergodicity was loosely defined and had no mathematical sense until the early 1930s, when first von Neumann and later, George D. Birkhoff reformulated and settled its mathematical basis [106]. Highly influenced by the contributions of Hilbert in the axiomatization of mathematics [111], they demonstrated the ergodic hypothesis that led to the most known Birkhoff theorem [112]. The ergodic hypothesis is key for many fields and specifically, for supporting the use of Molecular

Dynamics in the study of protein folding. Protein folding studies not only cover the conformation of protein backbone and its side chains (residues) when it is handling a given function, but also involves the study of the different protein conformations over time and the process that leads to its folding or unfolding [113] [114]. If the folding process was the result of trial and error method applied to all random possible movements of each atom in the protein backbone, protein folding would be eternal. However, it is not! Indeed, protein folds in a time scale of seconds or less. This apparent contradiction led to the formulation of the Levinthal's paradox [115]. The solution to this paradox is obvious: the energy cost of some movements by many possible agents (interatomic forces, protein-solvent system, etc.) drives the protein to a global free energy minimum [113]. Hence, by computational simulation it is possible to solve simultaneously the forces governing the motion of all atoms forming the protein, namely Newton forces (see the transcendence of Newton contributions), in order to obtain the different conformations that the protein adopts over time represented by its coordinates in a given reference at each time-step [116]. This methodology known as Molecular Dynamics (MD) has its beginnings in 1950s when first MD simulations of simple fluids were done. However, the breakthrough to understand protein folding was carried out in 1977, when McCammon et al. [117] published the first MD simulation of a protein based on the groundwork carried out some years before by Levitt, Lifson and Warshel among others [116] [118]. However, what is the point to carry out a time simulation if the real interest is the conformation of an ensemble system made up from several folded proteins in a single time? Here is where ergodicity comes to scene. As stated by Oostenbrink et al., the ergodic hypothesis demonstrates that *“one can simulate a single molecule with its surroundings for a period of time and get time-averaged molecular properties that approach the experimentally measurable ensemble averages. This means that from a*

*simulation of a system in time, we can get conformations that correspond to a thermodynamic ensemble or state point*" [116]. Thus, if most probable stable protein conformations of the MD over time are found, those will represent the majority of the population in an ensemble system of the protein, and not only this, if the given protein has any specific functionality, this will occur more likely when the protein adopts this conformation [118]. This property is very useful in many biological and natural science fields, such as discovering how enzymatic processes work, how protein binding occurs for drug design or how protein structure is related to its properties [118] [116] [119]. However, to find the most stable conformations the local and global minima of the full trajectory must be found by comparing the Gibbs free energy of each conformation with the rest. The equation that permits to calculate the Gibbs free energy in terms of the probability distribution of the ensemble is a direct consequence of statistical mechanics, since it is related to the Gibbs statistical interpretation of the entropy and related to Boltzmann's interpretation [120]. For the purpose of identifying those minima, Gibbs free energy of each conformation must be represented in a geometric surface, namely the Free Energy Landscape (FEL). In order to understand the folding process of a protein, its FEL must be deeply studied in such a way that minima and energy barriers that correspond to the transitions between minima can be characterized to provide an explanation of the fluctuations between the conformations represented between minima [120] [113]. In order to represent the FEL with a reduced dimensionality, without the loss of much information, original coordinates extracted from MD must be treated and expressed in a different basis. Therefore, computational methods to identify collective motions are usually applied, being one of the most commonly used the Principal Component Analysis (PCA) [121]. Originally introduced by Pearson in 1901 and later by Hotelling in 1933, PCA translates the original data to new variables in a different basis that successively

maximize variance and that are uncorrelated between them [122]. This has to be with finding the eigenvectors and the eigenvalues of the covariance matrix, after center original data. Those eigenvectors, known as principal components, conform the convertible matrix that applied to original centered data determine the new coordinates in this principal component basis. Besides this, eigenvalues are related to the percentage of covariance explained by its correspondent eigenvector, so the eigenvectors ordered by decreasing eigenvalues conform the ordered principal components [120] [121]. Daidone and Amadei state that “Typically, more than 90% of the total atomic fluctuation is described by  $\approx 20\%$  of the principal axes” [121]. Thus, FEL might be represented using two or three principal components and still save the information of the original data. In the fifth and final work presented in this thesis, a Molecular Dynamics simulation followed by a Principal Component Analysis, the building-up of the Free Energy Landscape and a posterior clustering analysis was carried out [123]. The main goal was the characterization of key residues relative position of Bombesin peptide, which might be isolated from the European frog *Bombina orientalis*, and that is involved in a wide spectrum of biological activities, such as satiety feeling, control of circadian rhythm, corporal thermoregulation or stimulation of gastrointestinal hormone release among others [124]. Nevertheless, the main interest about Bombesin is related to its affinity for the BB2 G protein-coupled receptor, its lower affinity for BB1 and its non-affinity for the BB3 receptor [125], which makes of Bombesin an interesting peptide for drug design, particularly in the fight against cancer. Similar long-sequence peptides were previously characterized like those reported by V. Erspamer and P. Melchiorri [126] or by R. T. Jensen and T. W. Moody [127]. For this fifth work though, two main references have been taken into account, previous spectroscopy studies including NMR, especially those in 2,2,2-trifluoroethanol/water mixture (30% v/v) by J. A. Carver and J. G. Collins [128]

and Díaz et al. [129], and previous computational studies addressed to assess the conformational profile of Bombesin by Sharma et al. [130]. Results agree with those reported by these references, showing the coexistence of two main conformances, a less likely N-terminus hairpin structure and a most probable C-terminus  $\alpha$ -helix conformation, being the side chains of Trp8 and His12 in this partial helix structure the residues responsible to adopt a conformation suitable for binding [123].

Finally yet importantly, this thesis as well as most of late XX century and this XXI century developments in physical sciences would not have been possible without those powerful and indispensable working tools at which everybody seem to be attached every day: the computers. The development of those machines is a contribution of many scientists over history (as all other fields discussed above, as it seems), whose origins stands in Leibniz's combinatorics, starting point of the electronic calculator, and Turing's logic machine [131]. Here, technology developments by Leibniz merged with Turing's conceptual mathematics in what is another example of how those branches have contributed to the progress of the other. Besides fundamental prematurely contributions as Charles Babbage's analytical engine and George Boole's algebra, von Neumann architecture represented a real milestone on computer science, as only after his collaboration with the ENIAC (Electronic Numerical Integrator And Computer, 1944-1955) team led by Herman H. Goldstine first results on joining highspeed, high-precision, automatic calculation and reasoning capabilities were achieved [131] [132]. Few years later von Neumann wrote the "*First draft of a report on the EDVAC*" presenting the Electronic Discrete Variable Automatic Computer (EDVAC) that, although using vacuum valves like ENIAC did, it substituted the decimal system of ENIAC by a binary one [133]. While von Neumann architecture sat the basis for modern computer hardware

architectures, first programming softwares were developed in the 1950's, such as COBOL, ALGOL, LISP or FORTRAN (developed by IBM) among others [131] [134]. Despite being the oldest ones, some of them are still widely used today in some fields. FORTRAN, for example, is the language used for programming the Molecular Dynamics and posterior analysis softwares in the fifth work presented in this thesis. More modern programming languages appeared in 1970s, such as C language developed by Bell Laboratories [135]. The impact of all these achievements in computer science might be easily seen in all of the fields discussed so long, as none of them could have been possible without computing machines.

It would be absolutely unfair to state that all contributions that relate mathematics and physical sciences discussed so far have been only man's achievements, as they could not have been possible without the contribution of brilliant women who settled the groundwork of most of them. Early mathematics development would not have been the same without Hypatia of Alexandria contributions, with numerous mathematical commentaries like those published in *Arithmetica* by Diophantus (discussed above as transcendent for algebra divulgation) or *On the conics of Apollonius* by Perga. In early 1800s Sophie Germain contributed to the demonstration of Fermat's Last Theorem and, despite not being successful, her work influenced other mathematicians (some of them so reputed as Gauss) on what would be called later number theory field. Newton and Leibnitz differential calculus owes much to Émilie du Châtelet translation and divulgation of their work, and Maria Gaetana Agnesi's work *Instituzioni* (1748), a discussion of differential and integral calculus that harmonized those fields and that contributed with the "Curve of Agnesi". The following century mathematicians and scientists were captivated by Mary Somerville, who achieved enormous popular success and influenced

many great scientists, such as Laplace or Maxwell. In late XIX century, Sofia Kovalevskaya contributed to the analytic partial differential equations field with the Cauchy-Kovalevskaya theorem that became a groundbreaking novelty in PDE analytical solution, with application to Fourier's heat equation and several other physical phenomenon equations (like those presented in the second and fourth work of this thesis). Later in the early XX century, Emmy Noether, a Göttingen colleague of Hilbert, proved her Noether's theorem, a mathematical demonstration stating that laws of physics are independent of time or place, which was essential for the foundation of quantum physics. Closely related to this field, Marie Curie-Skłodowska and her daughter Irène Joliot-Curie are well known for their study of radioactivity (for which both won a Nobel Prize in 1903 and 1935 respectively), having Marie and her husband Pierre discovered Polonium and Radium elements. The assistant of Marie Curie, Ellen Gleditsch, was responsible of other remarkable contributions, as half-life of radium and the existence of isotopes. Other notorious scientists in nuclear and atomic theory were Lise Meitner, with the discovery of nuclear fission that changed history with the beginning of the "Atomic age", and the second woman to win Nobel Prize Marie Goeppert-Mayer for her structure of the nuclear shell model. As far as statistical mechanics, at early 1900s both Tatyana Ehrenfest-Afanassyeva and her husband Paul Ehrenfest reviewed and contributed significantly to the work of Boltzmann, helping to the development of the statistical interpretation of mechanics toughly critiqued at the beginning. Probably, the fifth work in this thesis could not have been possible without those contributions to statistical mechanics, and neither without the achievements of Dorothy Crowfoot Hodgkin, Nobel Prize for chemistry in 1964, who dedicated her life to decipher protein and enzyme structures by X-ray crystallography (such as penicillin and vitamin B12), and Rosalind Franklin, who in 1950s discovered the double alpha-helix structure of DNA proteins. And what to say

about Ada Lovelace, the “Mother of Computers”, the first computer programmer who surpassing Babbage’s mere mechanical calculator demonstrated the real potential of computers in the first half of XIX century. Following Ada steps, between 1940s and 1970s Rózsa Péter broke new ground with recursive functions, a new mathematical field with impact in computer science. Those are only some examples extracted from *Woman in Science* by the European Commission [136] of how great women, besides having been excluded from the scientific realm for so long, made significant contributions to most of the fields discussed above, and without them relation between mathematics and physics would have been orphan.

So far, we have described examples of how mathematics has contributed significantly to the development of physical sciences or vice versa, how physical laws have forced mathematical breakthrough under demand in different fields. However, this thesis is specifically focused on the application of mathematical methods to chemical engineering problems, demonstrating how mathematical tools enable to obtain significant physical results. This long-standing relation is not new in chemical engineering either. From the use of geometric similarity for scaling laboratory experiments into industrial plants, to the use of algebra and differential calculus to solve diverse problems, they are taught in chemical engineering courses more than 120 years ago [137]. The real breakthrough on mathematics and chemical engineering relation came with mathematical modelling, discovering that this technique made simulation and prediction of real chemical systems behavior possible [138]. It was on early 1900s that first books entirely aimed at mathematical methods applied to chemical engineering appeared, namely 1902 book "*Higher Mathematics for Students of Chemistry and Physics with Special Reference to Practical Work*" by J.W. Mellor [139] and 1906 book "*Fourier Series and Integrals*



*and Mathematical Theory of the Conduction Heat*" by R.S. Carslaw [140]. Both books focus on the solution of differential equations raised from thermodynamic, kinetic or chemical equilibria problems in Mellor's book and heat transfer problems in Carslaw's one. In the 1940s, Neil Amundson became notable for his works on separation and chemical reaction processes modelling, acting as a reference for follower authors like Rutherford Aris, coworker of Amundson, who contributed with a dissertation on chemical reactors design [137]. In the 1970s and 1980s non-conventional solution methods for differential equations appeared, being the orthogonal collocation from Villadsen et al. [100] a highlight between them. Later, Bird, Stewart and Lightfoot published "*Transport Phenomena*" [141], where empirical transport phenomenon was modelled and solved in detail. From 1970s onwards, with the improvement of computational science, many new works about numerical methods applied to chemical engineering systems appeared, optimization methods applications, routines in FORTRAN language, and the age of software packages as ASPEN, DESIG, PROCESS and HYSIM started [137]. From 1960s also date the first "stochastic models" applied to chemical engineering, as Markov chains applied to the study of batch chemical reactors kinetics by Fredrickson (1966) [142] or residence time distributions by Fan et al. (1981) [143]. Following stochastic models, neural networks applications to chemical engineering problems became popular between 1990s and 2000s [137].

This overview of either how advances in mathematics have supposed a significant impact in physical sciences development and vice versa, how physical findings contributed to mathematical improvement, is where the purpose of this thesis has its foundations. Furthermore, each of the issues explored in the works of this thesis has been presented as a milestone of the mathematical and physical merging history, and as far as

chemical engineering is concerned, the structure of the thesis reflects the historical evolution of the main mathematical methods applied in chemical engineering research. Hereby, this thesis can also be seen as a compendium of the fundamental mathematical methods to which every chemical engineer should be concerned when models of most common chemical phenomena arise.

To conclude in Alexander Pope's words: "This mighty maze is not without a plan"  
[1]. Well, mathematics mostly enlightens those plans hidden in physical phenomena and, eventually, lead us to unravel this mighty maze in where we all live.

### 3. Results and discussion

The results of the present study are reported in the form of a compendium of five publications. They are analyzed from two different perspectives: on the one hand, the suitability of the mathematical methods utilized for their impact in understanding diverse chemical engineering problems and, on the other, those related to the novelty of the results obtained that permit a better description of the process involved.

Each of these works, besides approaching diverse chemical problems, uses different solving strategies aimed at the study of a different mathematical method, providing novel results for the embedded physical phenomena. Thus, a double goal is discussed: the suitability of the mathematical method applied and the coherence and consequences of the scientific contributions, whether it may be novel parameter values or a conformational analysis.

To this point, it seems obvious that both are related, being the physical and chemical results the best evidences of the mathematical methods influences in chemical engineering science.

The main physical and chemical phenomena discussed are chemical kinetics, thermodynamic properties of supercritical mixtures, supercritical extraction of solid-fluid and liquid-fluid systems, and peptide conformational assessment. While optimization, orthogonal collocation, neural network design, finite differences method, molecular dynamics and Principal Component Analysis conform the bulk of mathematical methods employed.

Concerning quality indicators of the impact of the publications it is very early, but it can be assessed a priori by the impact factor of the journals they are published. Among the five presented papers, two have been published in a 1st quartile journal, one has been accepted for publication in a 2nd quartile journal and the remaining two are under review by the time this thesis is submitted.

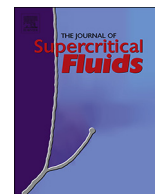
### **3.1. Binary interaction parameters from reacting mixture data. Supercritical biodiesel process with CO<sub>2</sub> as cosolvent.**

(Published in The Journal of Supercritical Fluids, 2019, 143, 107–119)

The reaction mechanism of oil triglycerides with methanol in presence of supercritical CO<sub>2</sub> for biodiesel production is a well known process in literature [12] [13] [14] [15] consisting of three reversible reaction in series. However, complexity raises when the evolution of the concentration of all species present in the reactor over time are intended to be modeled, as mixture not only depend on kinetic mechanism as also depends on the evolution of the thermodynamic state. As Binary Interaction Parameters (BIPs) of intermediate species is unknown for most of cubic Equations of State (EoS), a solving method mainly based on global optimization was designed to find the six kinetic parameters and the fourteen unknown BIPs of Peng-Robinson EoS by adjusting model to experimental values from group previous work [15]. Critical data from the mixture species, namely triolein, diolein, monolein, methanol, glycerol, FAME (Fatty Acid Methyl Ester) and CO<sub>2</sub>, was taken from literature (see Annex I – On critical properties from literature), along with seven values from the twenty-one different BIPs that those species determine (see Annex I – On BIPs literature values and estimation). Whether previous works had already determined BIPs from supercritical experimental data [26] [27], this work presents a groundbreaking method to determine them simultaneously with kinetic parameters for a biodiesel production model. Two main global optimization MATLAB software routines where considered during the resolution: *GlobalSearch* and *Genetic Algorithm (GA)* [49], showing the former a better performance, as both converged to similar results but with a much higher computational cost in GA case (see Annex I – On Optimization Discussion).

Direct kinetic parameters obtained were relatively well determined, agreeing those previously published [12]. Furthermore, a first direct kinetic parameter estimation could be extracted from initial reaction rate, namely the first slope of triolein concentration curve, showing a similar order of magnitude than simulation results. Inverse kinetic parameters, presenting more uncertainty with large confidence intervals though, revealed irreversibility in the last reaction. Energy activation from Arrhenius expression also agree with typical catalyzed processes [144] [145]. As far as BIPs results is concerned, values obtained at 150°C and 180°C were rejected because of the high error and lack of accuracy, while significant results were obtained at 205°C. Those BIPs presented same order of magnitude values, coherence related to similar pair of species, and those with unlike interaction, namely between polar and non-polar species, agreed in negative sign [146].

As seen, the above paper contributes with novel BIPs for Peng-Robinson EoS and a rigorous study of kinetic mechanism of biodiesel production from supercritical triolein trans-esterification in presence of methanol. Those results are the best evidence of the mathematical method goodness, presenting a powerful tool for future similar models solving. Hence, this work by Abel Valverde and Francesc Recasens was published in *The Journal of Supercritical Fluids* (2019), 1<sup>st</sup> quartile in Chemical Engineering area and 2<sup>nd</sup> quartile in Physical Chemistry area, IF (2018) 3,481, with two citations so far [11].



## Binary interaction parameters from reacting mixture data. Supercritical biodiesel process with CO<sub>2</sub> as cosolvent

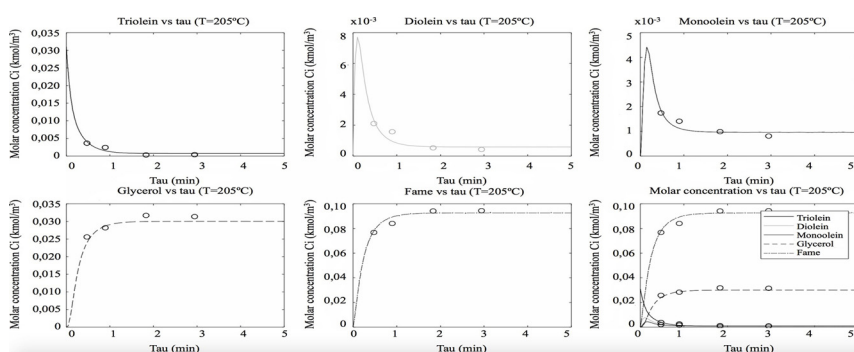
Abel Valverde<sup>a</sup>, Luigi Osmieri<sup>b</sup>, Francesc Recasens<sup>a,\*</sup>

<sup>a</sup> Department of Chemical Engineering, ETSEIB, UPC, Diagonal 647, E-08028 Barcelona, Spain

<sup>b</sup> National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA



### GRAPHICAL ABSTRACT



### ARTICLE INFO

**Keywords:**  
 Reactor model  
 Simulation  
 Peng-Robinson  
 Biodiesel  
 Methanol  
 Supercritical

### ABSTRACT

A method is presented for the calculation of binary interaction parameters from reacting mixture data. The process studied is the supercritical methanol transesterification of triolein with CO<sub>2</sub> as cosolvent. The reaction proceeds through 3 reversible reactions in series, on a solid catalyst (Nafion<sup>®</sup> SAC-13), for which the kinetics is known. To describe the reactor performance, the kinetic constants and the binary interaction parameters for the Peng-Robinson equation of state, are necessary, and were included in the reactor model. Previously published view-cell observations and thermodynamic calculations indicate that no condensation occurs in the reactor. Kinetic and thermodynamic parameters were determined by a global optimisation procedure. From the rate constants obtained, the process exhibits some reversibility, despite the excess of methanol used. Our methodology allows to obtain the binary interaction coefficients for the intermediate reaction species (dioleins and monooleins) that were unavailable so far. The method is sensitive to the errors in the measured concentrations.

### 1. Introduction

For the design of process equipment, it is necessary to know the thermodynamics of the fluid phases. For high-pressure processes, the thermodynamics are best represented by a suitable equation of state (EoS). In this approach, the binary systems data for the

thermodynamics of mixtures are used. Therefore, the knowledge of the pertinent binary interaction parameters (BIPs), appearing in the mixing rules, is essential. Often, for reactions proceeding in several steps in series, intermediate species cannot be synthesised for performing specific phase equilibrium studies. This is the case, for example, of the high-pressure transesterification of triglycerides with methanol.

\* Corresponding author.

E-mail address: [f.recabax@gmail.com](mailto:f.recabax@gmail.com) (F. Recasens).

<https://doi.org/10.1016/j.supflu.2018.05.023>

Received 19 April 2018; Received in revised form 21 May 2018; Accepted 24 May 2018

Available online 13 August 2018

0896-8446/ © 2018 Elsevier B.V. All rights reserved.

**Nomenclature**

a	Attractive constant, Pa.m <sup>6</sup> /mol <sup>2</sup>
A,B	Coefficients defined by Eq. (21)
a <sub>i</sub>	PR EoS, pure i coefficient, Eq. (17)
a <sub>ij</sub>	Binary interaction coefficient, = (1-k <sub>ij</sub> )(a <sub>i</sub> .a <sub>j</sub> ) <sup>0.5</sup>
b	Co-volume, m <sup>3</sup> /mol
b <sub>i</sub>	b coefficient for pure i, Eq. (18)
C <sub>i</sub>	Concentration component i, mol/m <sup>3</sup>
D <sub>ax</sub>	Axial Dispersion coefficient, m <sup>2</sup> /s
d <sub>p</sub>	Particle diameter, m
E <sub>i</sub>	Activation energy for k <sub>i</sub> , kJ/mol
F	Total molar flowrate, mols/s
F <sub>i</sub>	Molar flowrate, mols i/s
FO	Objective function, Eq. (23)
k <sub>i</sub>	Rate constant, i = 1–6, m <sup>6</sup> /(mol.kg.s)
k <sub>ij</sub>	Binary interaction parameter,
k <sub>ij</sub> <sup>n</sup>	Regression coefficients, Eq. (22)
L	Reactor length, m
M0	Average molecular mass, kg/mol
P	Pressure, Pa
Pé	Péclet number for mass, Eq. (6)
R	Ideal-gas constant, = 8.314 J/mol.K
Re	Reynolds number, Eq. (6)
r <sub>0</sub>	Initial rate, Eq. (25), mol/ kg s
r <sub>i</sub>	Reaction rate of i, mol/(kg.s)
r	Modified rates, Eq. (11), m <sup>3</sup> /(mol.kg.s)
r <sub>ij</sub>	Reaction rate of i in reaction j, mol/kg.s
S	Reactor cross section, m <sup>2</sup>
SV	Space velocity, 1/s

T	Absolute temperatura, K
u	Superficial velocity, m/s
W	Catalyst mass, kg
w	Total mass flowrate, kg/s
y <sub>i</sub>	Mole fraction component i,
Z	Compressibility factor,

*Greek symbols*

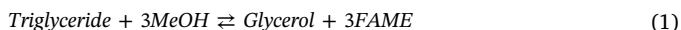
v <sub>ij</sub>	Stoichiometric coefficients, Eq. (7)
Ψ	Mass transfer group, Eq. (11), mol.kg.s/m <sup>7</sup>
τ	Space-time, s
χ	Peng-Robinson coeff. Eq. (19)
ω	Acentric factor,
ε	Porosity,
ν	Dynamic viscosity, m <sup>2</sup> /s

*Acronyms and subscripts*

BIP	Binary interaction parameter
BPL	Bubble-point line
CP	Critical point
DPL	Dew-point line
EoS	Equation of state
SC	Supercritical
FAME	Fatty acid methyl ester
SV	Space velocity
c	Critical
r	Reduced

In supercritical (SC) processes, cubic equations of state of the Van der Waals type (with appropriate mixing rules) are extensively used today to calculate component fugacities or concentrations in the reacting phases. They are mostly used for the prediction of vapour-liquid (VL) and liquid-liquid (LL) distribution of reactants and products in multiphase reactors. An example of a process with growing industrial interest is the SC transesterification of vegetable triglycerides with methanol to produce biodiesel, in the presence of an excess of CO<sub>2</sub>. In this process, conversions close to 99% are possible at very small reaction times [1–3]. In connection to this process, we have recently shown [4] that the density of the reacting mixture directly affects the concentration of the species within the reactor, as the density depends on the BIPs of the different species. Since mixture thermodynamics affects the kinetic model through gas-phase concentrations, the total number of parameters involved becomes generally large because many components are present.

The overall stoichiometry of biodiesel production, is as follows:



Where 3 mol of fatty acid methyl ester (FAME or biodiesel) are formed by transesterification (ester interchange) of triglyceride oil with 3 mol of methanol. Process (1) takes place stepwise through three reversible reactions in series, each one liberating an ester molecule, with the glycerol molecule generated in the last step, as illustrated by the following reactions:



Where A = triglyceride, B = methanol, D = diglycerides,

M = monoglycerides, E = fatty acid methyl ester (FAME, methyl oleate in our case), and G = glycerol. In our process [5], inert carbon dioxide, CO<sub>2</sub>, is added to the feed in order to lower its critical point. Integral reaction runs were made at 200 bar and temperatures from 150 °C to 205 °C on Nafion SAC13, a superacid ion-exchange solid catalyst [5], in a continuous reactor. Historically, the biodiesel processes were carried out at various conditions, mostly in liquid-phase batch reactors [6]. Operating in gas phase conditions, a continuous process in a tubular reactor becomes possible, in contrast with the liquid-phase, low tonnage, where batch operation is usually employed. Other ester interchange reagents (i.e., methyl acetate) have been also used in SC processing for biodiesel production [7,8], which offers certain advantages, as it is free from glycerol formation. Simulation studies on biodiesel manufacturing using methanol in SC CO<sub>2</sub> have been made [9,10]. Homogeneous or heterogeneous catalysts, or even no catalyst, can be used depending on the type of reactor. However, using a solid catalyst provides obvious advantages compared with a homogeneously catalysed process conducted in CSTRs, as in the case of the Henkel process [6], provided that catalyst deactivation is tolerable. The advantage of heterogeneous catalysis on the energy usage for SC process conditions has been pointed out as well.

In our previous articles [5,11] we used the solid catalyst Nafion® SAC -13, in a bench-scale, packed-bed reactor, since a continuous plug-flow reactor process would be the preferred industrial choice. Manufacturer's data for the SAC catalyst show its superior activity compared to the standard ion-exchange catalysts (Amberlyst-15, Dowex 50) in a number of liquid-phase organic reaction types. An ASPEN simulation study of a SC process was presented by Glisic and Skala [10], where they compared the energy input for different process alternatives (subcritical vs. super-critical process, homogeneous vs. heterogeneous catalysis, long-time catalyst activity vs. catalyst costs, etc.). They concluded that heterogeneous catalyst could be a good choice in view of reduced methanol excess. Recently, Kolb and co-workers [3] presented



a demonstration plant for SC continuous biodiesel production using  $\text{La}_2\text{O}_3$  on  $\gamma\text{-Al}_2\text{O}_3$  as catalyst placed in a specially designed micro reactor assembly. Very large space-velocities were possible with an otherwise stable catalyst. However, for their extreme process conditions (about 400 °C, 400 bar), free-fatty acids are formed. Recently, we proposed a novel idea based on a two-reactor SC process, which we modelled on the basis of the experimental kinetics available on Nafion SAC-13 [11]. In our two-reactor process [5], an inter-stage selective glycerol removal (effected by partial expansion and recompression of the fluid) was found to increase reaction rate and triglyceride conversion by shifting chemical equilibrium to the formation of products, resulting in reduced space-time values. Our work shows that much can yet be done to improve the SC biodiesel process, including the kinetic modelling to account for free-fatty acid formation [3].

The use of cubic EoS of the van der Waals type has been found most successfully employed in SC fluid processes [14–18]. The Peng-Robinson [19] modification of the Redlich-Kwong EoS [20] is used extensively today, as summarized by Reid [21]. The EoS for a pure fluid can be extended to mixtures via the mixing rules, containing one BIP,  $k_{12}$ , for a binary mixture, with the quadratic mixing rules for the  $a$ -value [22]. The application of the EoS to multicomponent mixtures involves the consideration of all possible  $i$ - $j$  pairs present. Brunner and co-workers [23], have developed a software (PE 2000), which considers more than fifty EoS, with several mixing rules for the most popular ones. In general, despite several attempts to correlate BIP have been made [14,24–27], the BIPs are still measured in phase equilibrium experiments. In the predictive method of Jaubert et al. [27], the database is available today for up to 40 chemical groups and gases, related mostly to petroleum fluids using the predictive PR-EoS [28]. Other GC-EoS methods have been found appropriate for non-reactive and reactive SC fluid engineering studies [28,29].

In this work, we present a simulation model for the SC biodiesel production process. For high-pressure process simulation, the approach used by the most common simulators (ASPEN, CHEMCAD, ProSim, PROII, etc.) to predict VL phase equilibria is based on the EoS method (i.e., the phi-phi approach). In this case, two parameters are calculated, that is, the  $a_i$  ( $T$ ,  $T_c$ ,  $P_c$ ,  $\omega$ ) and the  $b_i$  ( $T$ ,  $P_c$ ,  $T_c$ ), where  $i$  are the pure species,  $T_c$  is the critical temperature,  $P_c$  is the critical pressure, and  $\omega$  is the acentric factor. In the case of the reactions of Eq. (2), there are 6

reactants and products, plus inert  $\text{CO}_2$  added on purpose. Thus, 7 species are present in total. Then, the mixing rules are applied to each binary system. So, in principle, the number of possible BIPs for 7 components becomes quite large. Furthermore, to model the reactor, the rate constants for the different reactions are necessary. In a high-pressure vapour-phase process, component fugacities, activities, or molar concentrations are required to express the reaction rates. So, the kinetics of the above reactions, Eq. (2), requires 6 more kinetic constants,  $k_i$  ( $T$ ), and their activation energies. Then, the conservation equations for the reactor can be written. On the other hand, if the reacting fluid composition is available from experiment in a given reactor at different space-time and temperature, the BIPs,  $k_{ij}$ , and the rate coefficients,  $k_i$ , must be obtained simultaneously by fitting the measured concentrations in the fluid of the reactor. This is the case studied in the present work.

To the best of our knowledge, only two research groups have attempted to determine the BIPs using data on a SC reacting mixture. The first group was that of Bertucco et al. [30]. Their process was the SC hydrogenation of a pharmaceutical organic intermediate carried out in a trickle-bed reactor. The authors directly calculated the  $k_{ij}$  values (using PR-EoS) to fit the observed phase distribution of reactants in the gas and liquid phases. A second research group [31], used a rather indirect calculation to obtain the  $k_{ij}$  (reaction was the production of the synthesis gas). In their approach, they first used UNIFAC, and then they combined the UNIFAC results with the Redlich-Kwong-Soave-EoS to obtain by regression the binary parameters of the multicomponent mixture. In the present work, we have used a direct approach similar to that of Bertucco et al. Thus, our method is based on BIPs optimisation to fit the observed concentrations of reactants, products, and inerts, at the exit of a gas-phase reactor.

Our purpose in this work was first to propose a novel methodology for the joint determination of binary parameters,  $k_{ij}$ , and rate constants,  $k_i$ , for the case where fluid properties are described with the PR-EoS. Secondly, we applied this method to develop our key tool, i.e., a model for a continuous, integral catalytic, packed-bed reactor for the production of biodiesel from triolein and methanol with  $\text{CO}_2$  as cosolvent. To this end, the measured concentrations of reactants, intermediate species, and products for a bench-scale reactor, were analysed with our reactor model that uses available catalytic rate equations from a

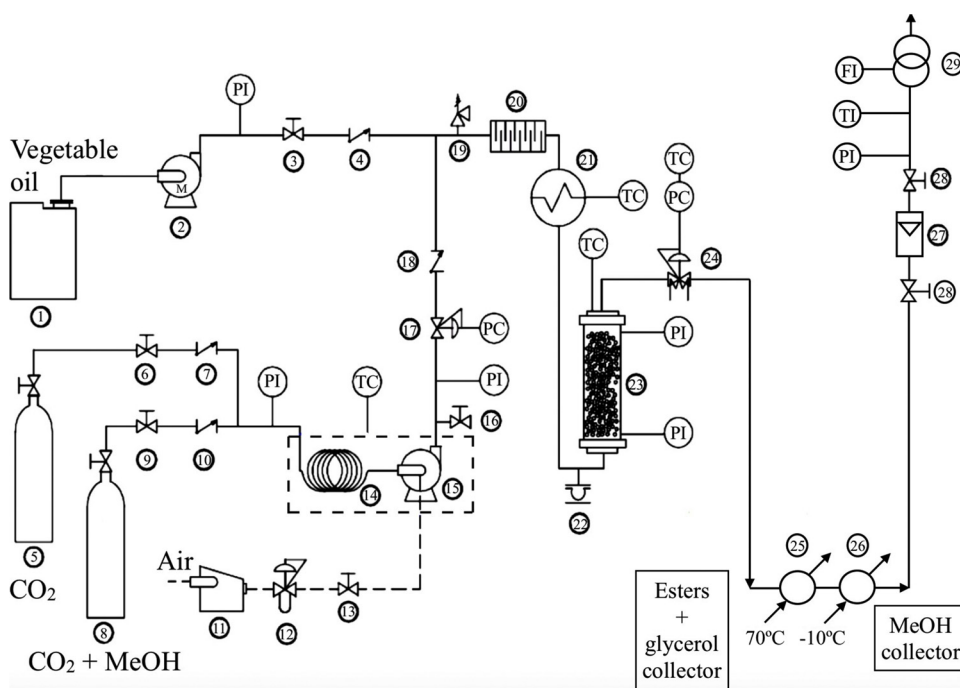


Fig. 1. Experimental setup. 1-Oil reservoir; 2-Pump; 3,6,9,13,16,28-Ball valves; 4,7,10-Check valves; 5- $\text{CO}_2$  Cylinder; 8- $\text{CO}_2/\text{MeOH}$  mixture cylinder;11-Air-driven compressor; 12- Pressure regulator;14- Cooler; 19-Relief valve; 20-Static mixer; 21-Pre-heater;22-Rupture disk; 23- Reactor; 24-Backpressure valve; 25- High temperature bath; 26- Low temperature bath; 27-Rotameter; 29-Gas-meter.

previous work. As will be shown, a large number of different BIPs is required for the thermodynamic description of the vapour-phase of the reacting mixture containing 7 different components. Some of the BIPs were available from existing phase-equilibrium studies, but most of the BIPs for the reaction intermediates (monoglycerides, diglycerides, glycerol) are not available. Thus, the unknown parameters have been calculated by a global optimisation method.

## 2. Experimental section

### 2.1. Background

The experimental study behind this work was presented elsewhere [11]. Here we report only the most significant features. The detailed process flow diagram, showing feed system, instruments and controls is shown on Fig. 1. A summary of the experimental reactor and the catalyst used is given in Tables 1 and 2.

The plant was operated at steady state conditions. Note that in our set-up, a continuous fractionation of the reactor outlet stream was carried out in two heat exchangers in series (constant-temperature baths) operated at atmospheric pressure. The first bath was operated at 70 °C where the heavier components (triolein, FAME, and glycerol) were condensed. The remaining effluent gas (CO<sub>2</sub> and methanol vapours) was cooled in a second bath kept at -10 °C where methanol was condensed. Finally, the remaining gas stream (only CO<sub>2</sub>) was measured in a dry-gas meter. For detailed operating procedures and chromatographic methods, refer to our previous publication [11].

### 2.2. Insuring gas-phase conditions in experimental reactor

Two types of checks were made to make sure that the reactor was operated under gas phase conditions. The first check is based on thermodynamic calculations [5]. The second check is based on direct experimental observations using a view-cell [13].

#### 2.2.1. Thermodynamic calculations

The first check was based on the calculations [5] using the RK-Aspen EoS. Specifically, pure component properties (molecular weight (MW), boiling point (T<sub>b</sub>), T<sub>c</sub>, P<sub>c</sub> and ω), were taken from the ASPEN database. For the critical constants, the recommended values were used [9]. Where necessary, recognised methods of estimation [22,32] were employed. The values taken from the ASPEN database are given in Table 3. As discussed in the Introduction, we proposed [5] a two-step process with expansion and recompression of the flowing stream in between the two reactors. In our process [5], the use of the PT diagram was necessary for avoiding condensation in the reactor. The PT diagram is shown in Fig. 2 [5]. The bubble-point line (BPL) and dew-point line (DPL) were calculated using the ASPEN Plus process simulator. The critical point and the operating P and T for the fluid in the first reactor are marked on the figure. As seen on Fig. 2, condensation can be safely ruled out.

#### 2.2.2. View-cell observations

Recently some authors [13] made view-cell observations that lent credibility to the previously described predictions [5]. These authors performed isothermal compression of the reacting mixture to different final pressures in a view-cell, and their PT data were then plotted as experimental data points on a PT diagram. These points, located on both sides of the critical point of the mixture, fell between the BPL and DPL. The observations were done on a mixture corresponding to the reactor inlet conditions. In the reactions of Eq (2), it is evident that the products are more volatile than the reactants. Therefore, if the reactant mixture is seen in the view-cell to be a homogeneous gas-phase [13], with more reason the mixture of reactor outlet will remain homogeneous. Other authors have published view-cell observations for a similar SC process. The method of Tavlarides and co-workers [33] is

recommended if future researchers perform view-cell observations (Fig. 3).

## 3. Theory

First, it is important to have a reliable reactor model that can be used to predict performance for changes in feed composition, operating temperature and space-time. In this section we present the analysis of plug-flow behaviour, with reaction kinetics, and a simulation model for the reactor. Our second purpose here is to define the equation of state used to calculate the thermodynamic features of fluids.

### 3.1. Reactor feed and output streams

The bench-scale plant (Fig. 1) was operated continuously at steady-state conditions. As seen on Table 2, four different flow rates and three temperatures were tested in the reactor. After changes in process settings (flow rate and temperature), new steady-state conditions were ensured by waiting sufficient time before analysing the reactor effluent. Analytical procedures to measure the reactor feed reactants and output product concentrations were described in detail elsewhere [11].

### 3.2. Reactor considerations

We first justify that plug-flow conditions prevailed in the reactor (see Fig. 1 and Table 1), as this is a key assumption for the model. Note that volume-equivalent particle diameter for the catalyst was given in Table 1. Moreover, the bed void fraction was around 0.75. To assess axial dispersion in the experiments, we need to calculate the Reynolds and Péclet numbers for flow in the packed bed. These are defined as:

$$Re = \frac{u d_p}{\nu} \quad Pe = \frac{uL}{D_{ax}} \quad (3)$$

Very few authors have studied dispersion for flow of SC fluids

**Table 1**  
Properties of reactor and catalyst.

Type	Tubular type (Titanium) <sup>a</sup>
<b>Reactor</b>	
Cylinder inside dimensions	152 mm x 15.5 mm
Volume	26.8 cm <sup>3</sup>
Cross section	188.69 mm <sup>2</sup>
Heating resistance	2.5 kW
<b>Catalyst<sup>b</sup></b>	
Designation <sup>b</sup>	Nafion <sup>®</sup> SAC-13
Catalyst loading. W	9 g
Type	SAC/amorf. SiO <sub>2</sub> composite
Nafion content	13%
Extrudates	1 × 9.4 mm
Pore volume. V <sub>g</sub>	> 0.6 cm <sup>3</sup> /g
Surface area. S <sub>g</sub>	> 200 m <sup>2</sup> /g
Acidity type	Brønsted only
Active sites conc.	0.13 meq/g
Active groups	-CF <sub>2</sub> -CF <sub>2</sub> -SO <sub>3</sub> H
Nafion clusters size	20-60 nm
Moisture	1-3%
<b>Other catalyst properties</b>	
Internal porosity. ε <sub>p</sub>	0.57
Pore diameter, calculated <sup>d</sup>	12 nm (micropor.)
Particle diameter. d <sub>p</sub>	2.416 mm
Bulk density, measured	0.314 g/cm <sup>3</sup>
SAC-13 vs. pure Nafion activity <sup>c</sup>	118.7 times larger
Hammet value	-14 to -10

<sup>a</sup> Eurotechnica (Germany).

<sup>b</sup> DuPont-Engelhard (USA) [12].

<sup>c</sup> Dodecene-1 liquid phase isomerization rate at 70 °C.

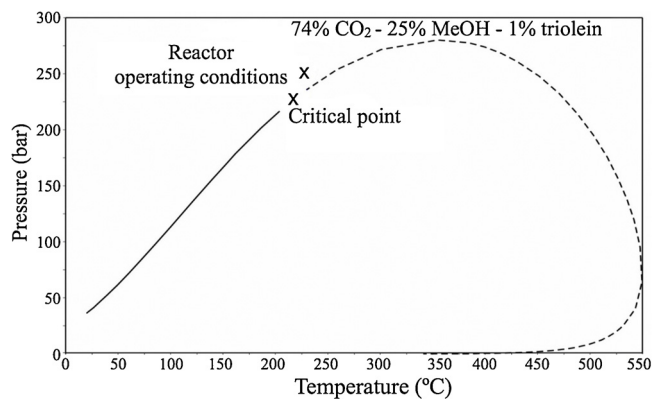
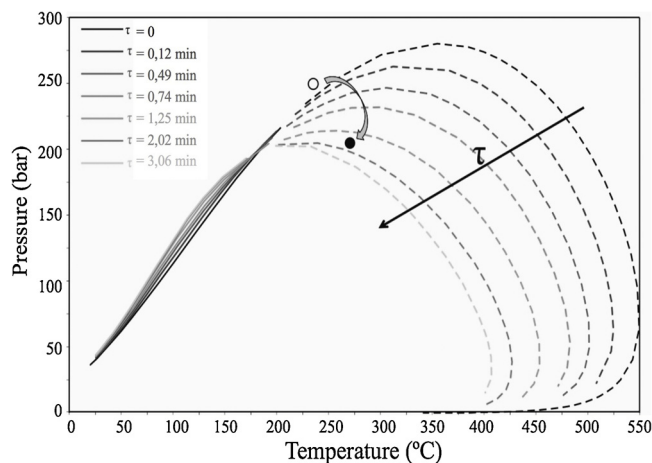
<sup>d</sup> Farrauto and Bartholomew [34].

**Table 2**  
Experimental conditions.

Temperature	150°, 180°, 205 °C
Pressure	250 bar
Space-time, $\tau$	0.5–4 min
Vegetable oil type	Sunflower
Oil mass flow rates	0.2–1.4 g/min
Solvent mass flow rate	16.2 to 2.3 g/min
Feed composition, mass %	CO <sub>2</sub> /Me OH/Oil, 69.4/23/7.6

**Table 3**  
Physical Properties, Critical Constants and Acentric Factors of Components (ASPEN database).

	Triolein	Dirolein	Monoolein	Methyl oleate
Molecular weight (kg/kmol)	885	621	356	296
Normal Boiling, Tb (K)	827	765.03	674.82	596
Liq. volume at Tb (m <sup>3</sup> /kmol)	2.71	1.11	0.53	0.49
Critical temp., Tc (K)	977	920	835	721
Critical pressure (MPa), Pc	0.334	0.505	1.056	1.103
Critical volume (m <sup>3</sup> /kmol), Vc	3.25	2.83	1.254	1.108
Pitzer's acentric factor, $\omega$	1.97	1.76	1.53	1.04

**Fig. 2.** PT diagram for the feed mixture to the reactor [5]. Composition is in mole %.**Fig. 3.** Recession of the DPL for increasing space-times [5]. The multi-component mixture is that prevailing at reactor exit. (The white and black dots indicate a possible PT progress to avoid condensation).

[35,36]. We used axial mass dispersion information for flow of SC fluids [36] and represented their data on the Levenspiel and Bischoff graph [37]. The dispersion data of Abaroudi et al. [36] are for neat and modified SC carbon dioxide. In our reaction runs, Reynolds number was 7–80, so the values of the dispersion number can be taken as 0.6–0.7 from the dispersion graph [36]. From the values of bed void fraction and reactor length, the Péclet values are 152–180, hence large enough for plug flow conditions to be assumed ( $Pé > 100$ ).

### 3.3. Kinetics

For the reactions of Eq. (2), the following empirical catalytic rate equations given by Eq. (4) have been used. The kinetic model was successfully used in recent articles for the transesterification process [5,11,13,33]. This is:

$$\begin{aligned} r_1 &= k_1 C_A C_B - k_2 C_D C_E \\ r_2 &= k_3 C_D C_B - k_4 C_M C_E \\ r_3 &= k_5 C_M C_B - k_6 C_G C_E \end{aligned} \quad (4)$$

Mechanistic catalytic rate equations of the Langmuir-Hinshelwood, Eley-Rideal and other types [38], have been proposed [39] for this process for low-pressure, liquid-phase conditions on a different catalyst and feed composition. Unfortunately, similar kinetic information is not available for the high-pressure gas-phase process on our SAC-13 catalyst. In this paper, we assume that the kinetics of the 3 reactions of Eq. (2) is given by Eq. (4).

The relationship between molar concentrations and mole fractions for a high-pressure gas is as follows:

$$C_i = \frac{P y_i}{ZRT} \quad i = A, B, M, D, E, G, CO_2 \quad (5)$$

Where, Z, is the compressibility factor of the gas mixture, calculated from the PR-EoS (see section 3.5). The dependence of  $k_i$  on temperature is given by the following Arrhenius equation:

$$k_i(T) = k_{i0} \exp\left(\frac{-E_i}{RT}\right) \quad (6)$$

For the kinetic constants appearing in Eq. (4), for  $i = 1-6$ .

### 3.4. Catalytic reactor model

In this section we develop a simulation model in terms of the space-time (the reciprocal of space-velocity), a customary industrial measure of contact time in catalysis terminology. We will consider a one-dimensional, isothermal packed-bed reactor, where a slice of catalyst mass  $dW$ , occupies a differential volume,  $dV$ , of length  $dL$ . In view of the small bed length, pressure drop is assumed to be negligible. A steady-state mass balance for components in terms of the catalyst mass and reaction rates in the fluid phase, can be written as:

$$\frac{dF_i}{dW} = \sum_{j=1}^3 \nu_{ij} r_{ij} \quad i = A, B, M, D, E, G, CO_2 \quad (7)$$

Where  $F_i$  are the molar flow rates of species  $i$ ;  $r_{ij}$  is the net formation rate for species  $i$  in reaction  $j$  (see Eq. 4) and  $\nu_{ij}$  are the stoichiometric coefficients (all equal to  $\pm 1$  in our case). For inert CO<sub>2</sub>,

$$\frac{dF_{CO_2}}{dW} = 0 \quad (8)$$

The initial conditions for Eqs. 7 and 8, for  $W = 0$  ( $L = 0$ ) are given in Table 4 expressed in moles. In the above equations the catalyst mass is expressed in terms of catalyst bed volume using the bed density,  $\rho_B$ , (see Table 1). The above mass balance system can be written, as:

$$\frac{dF_i}{\rho_B S dL} = \frac{F dy_i}{\rho_B S dL} = \sum_{j=1}^3 \nu_{ij} r_{ij} \quad (9)$$

Where F is the total molar flow rate in the reactor cross-section, S; and L is the reactor length coordinate. F is the sum of all reacting species

**Table 4**  
Initial Mole Fractions for Reactor Model, L = 0 (P = 250 bar).

Temperature	Me OH $y_{Bo}$	Triglyceride $y_{Ao}$	Carbon dioxide $y_{(CO_2)o}$
150 °C	0,3129	0,0041	0,6831
180 °C	0,3129	0,0036	0,6835
205 °C	0,3129	0,0037	0,6834

All product mole fractions at bed inlet are zero.

including inerts. Pressure is assumed to be constant as reasoned before. As derived from the overall reaction stoichiometry, see Eq. (1), fluid density is constant for constant P and T within the reactor. Thus, in a given run, the parameter F is constant. Introducing the rates given by Eq. (4), as a function of concentrations, into the mass balance Eq. (9), and using mole fractions,  $y_i$ , by use of Eq. (5), the isothermal reactor model consists in the following set of initial-value ordinary differential equations:

**Table 5**  
Minima of the objective function<sup>a</sup> with Genetic algorithm and Globalsearch.

T, °C	FO <sub>1</sub> with GA	FO <sub>1</sub> with Globalsearch
150	1.2248	1.2253
180	0.2339	0.2339
205	0.2417	0.2549

<sup>a</sup> OF<sub>1</sub> after 3 iterations, see Fig. 4.

$$\begin{aligned} \frac{dy_A}{dL} &= -\Psi r_1' \\ \frac{dy_B}{dL} &= -\Psi (r_1' + r_2' + r_3') \\ \frac{dy_D}{dL} &= \Psi (r_1' - r_2') \\ \frac{dy_M}{dL} &= \Psi (r_2' - r_3') \\ \frac{dy_G}{dL} &= \Psi (r_3') \\ \frac{dy_E}{dL} &= \Psi (r_1' + r_2' + r_3') \\ \frac{dy_{CO_2}}{dL} &= 0 \end{aligned} \tag{10}$$

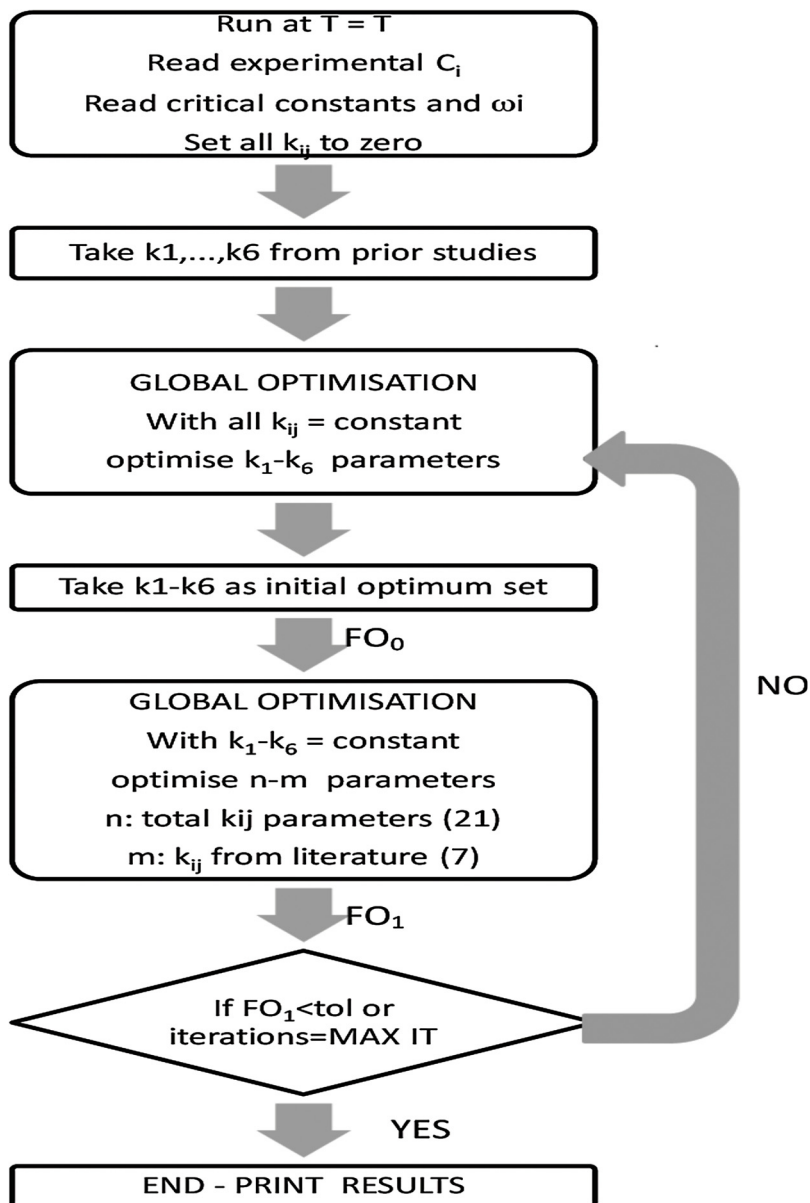


Fig. 4. Optimisation methodology for calculating the rate constants ( $k_i$ ) and the binary interaction parameters ( $k_{ij}$ ) for a run at temperature T.

**Table 6**  
Inlet mass flow-rates to reactor [11].

Temperature (°C)	CO <sub>2</sub> (g/min)	Methanol (g/min)	Triolein (g/min)	Total mass flow (g/min)
150	11.71	3.90	1.42	17.03
	5.90	1.97	0.72	8.59
	3.08	1.03	0.37	4.49
	1.71	0.57	0.21	2.48
180	13.43	4.48	1.42	19.33
	7.14	2.38	0.76	10.28
	3.53	1.18	0.37	5.08
	2.20	0.73	0.23	3.16
205	12.91	4.30	1.42	18.64
	6.90	2.30	0.76	9.96
	3.39	1.13	0.37	4.89
	2.11	0.70	0.23	3.05

Where the following symbols are defined in the Nomenclature section:

$$\begin{aligned}
 r'_1 &= k_1 y_A y_B - k_2 y_D y_E \\
 r'_2 &= k_3 y_D y_B - k_4 y_M y_E \\
 r'_3 &= k_5 y_M y_B - k_6 y_G y_E \\
 \Psi &= \frac{p_B SP^2}{F(ZRT)^2} \tag{11}
 \end{aligned}$$

The initial conditions for the system of equations Eqs. (9) to (11) are given in Table 4.

3.4.1. Reactor simulation

The integration of the reactor equations for a catalyst mass in the reactor, W, provides the mole fractions, y<sub>i</sub>, at any length coordinate, L, including the reactor exit (L<sub>max</sub> = 152 mm). For the integration of the model, a 4<sup>th</sup>-order Runge-Kutta algorithm was used. Two important points regarding the model equations deserve comment.

First, in order to calculate the gas-phase concentrations, a value of Z for the gas mixture at a location, L, in the bed is needed. The value of Z is calculated from the PREoS, using the VdW mixing rules. A value of Z involves the following calculations: the current mole fractions y<sub>i</sub>, the values for critical constants (T<sub>c</sub>, P<sub>c</sub>), the acentric factor for the pure species, and the BIP for each binary pair in the mixture. The critical properties are taken mainly from the literature, or estimated (see Table 3). In this way, for the 3 component pairs present at reactor inlet (i.e., Me OH-triolein, MeOH – CO<sub>2</sub>, triolein – CO<sub>2</sub>), a value of an initial Z can always be calculated. As will be discussed, in the first iteration the other BIPs are set to zero.

Second, note that the above mass balance equations constitute a useful simulation tool for a reactor holding a given amount of catalyst (W = 9 g in our case). Recall that a simulation model is not the same as a design model, although the conservation equations look the same. In the simulation model, the variable parameter for simulating the reactor is the molar flow rate, F, which changes from run to run for the various

**Table 7**  
Mass flow-rates measured at reactor outlet [11].

T (°C)	Methanol (g/min)	Triolein (g/min)	Diolein (g/min)	Monoolein (g/min)	Glycerol (g/min)	FAME (g/min)	CO <sub>2</sub> (g/min)
150	3.8926	1.2107	0.0958	0.0029	0.0043	0.1072	11.7126
	1.9594	0.5557	0.0616	0.0043	0.0029	0.0812	5.9045
	1.0180	0.2054	0.0372	0.0023	0.0035	0.0924	3.0838
	0.5502	0.0162	0.0065	0.0000	0.0063	0.1740	1.7069
180	4.4277	0.7500	0.1277	0.0250	0.0954	0.4638	13.4336
	2.3296	0.2948	0.0469	0.0097	0.0566	0.4707	7.1414
	1.1465	0.0680	0.0190	0.0052	0.0319	0.2798	3.5301
	0.7092	0.0128	0.0059	0.0031	0.0232	0.2076	2.1950
205	4.1828	0.1574	0.0647	0.0308	0.1170	1.1302	12.9149
	2.2293	0.0553	0.0257	0.0133	0.0688	0.6610	6.9022
	1.0908	0.0026	0.0041	0.0045	0.0380	0.3632	3.3901
	0.6791	0.0023	0.0021	0.0023	0.0234	0.2268	2.1110

flow rates (of Table 1) with fixed W. The relationship between molar flow rate and space-time is as follows:

$$\tau = \frac{W}{w} = \frac{W}{FM_0} \tag{12}$$

Where w is the total mass flow rate, that it is indeed constant along reactor length, L; M<sub>0</sub> is the average molecular mass of the inlet fluid gas mixture, and τ is the space-time (i.e., reciprocal of the space-velocity). In this way, the mole fractions y<sub>i</sub> are solved for each molar flow rate F. In summary, since F is inversely proportional to the space-time, the simulation model allows the prediction of mole fractions y<sub>i</sub> as a function of τ, by running the program at different flow rates, F. As it is common in continuous catalytic reactors, a single parameter is used to define the reactor contact time. This is the space-time, τ (or the space-velocity, SV). For a given reactor temperature T, by changing τ, the composition of the fluid at reactor output is changed, since contact time with the catalyst is changed.

3.5. The Peng-Robinson EoS

The pressure-explicit Peng-Robinson EoS is utilized in this work [19]. For a pure substance, the equation is:

$$P = \frac{RT}{v - b} - \frac{a}{v(v + b) + b(v - b)} \tag{13}$$

Where v is the molar volume, b is the excluded volume, and a is the attractive pressure constant. For the case of mixtures, a and b depend on temperature and composition. Assuming the one-fluid Van der Waals mixing rules for a and b, the following expressions for a, and b, of the mixture, were employed:

$$a = \sum_i \sum_j y_i y_j a_{ij} \tag{14}$$

$$b = \sum_i y_i b_i \tag{15}$$

The a<sub>ij</sub> is related to the geometric mean,  $\sqrt{a_i a_j}$  and the binary interaction coefficients, k<sub>ij</sub>, as follows:

$$\begin{aligned}
 a_{ij} &= \sqrt{a_i a_j} (1 - k_{ij}) \\
 k_{ij} &= k_{ji} \\
 k_{ii} &= 0
 \end{aligned} \tag{16}$$

The a<sub>i</sub>, b<sub>i</sub> for pure components are:

$$a_i = 0.45724 \frac{R^2 T_{ci}^2}{P_{ci}} [1 + \chi_i (1 - T_r^{0.5})]^2 \tag{17}$$

$$b_i = 0.07780 \frac{RT_{ci}}{P_{ci}} \tag{18}$$

$$\chi_i = 0.37464 + 1.54226\omega_i - 0.2699\omega_i^2 \tag{19}$$

**Table 8**  
SC Biodiesel Process: Final Values of Kinetic Constants with their Confidence Intervals.

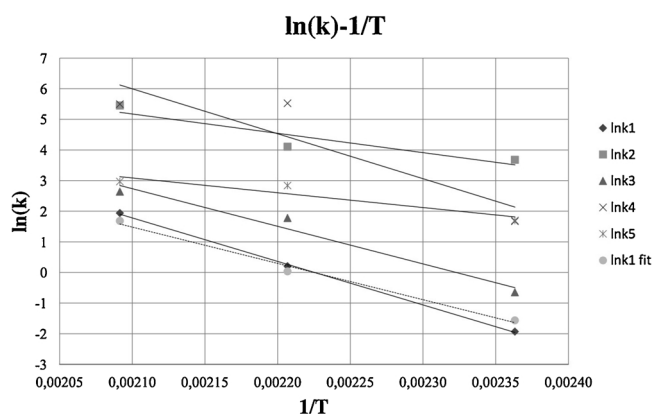
T (°C)	k <sub>1</sub>	k <sub>2</sub>	k <sub>3</sub>	k <sub>4</sub>	k <sub>5</sub>	k <sub>6</sub>
150	0.145 ± 0.07	39.57 ± 68.95	0.523 ± 0.46	5.326 ± 574.14	5.465 ± 3.1	0 ± 11.28
180	1.228 ± 0.3	60.650 ± 71.48	5.914 ± 1.8	250.508 ± 205.12	17.069 ± 4.3	20.981 ± 8.8
205	6.955 ± 22.26	282.83 ± 1238.0	13.996 ± 11.3	243.301 ± 172.3	19.476 ± 9.5	17.548 ± 10.4

Values are expressed in L<sup>2</sup>/(mol·kg<sub>cat</sub>·min). Multiply by 17 × 10<sup>-9</sup> to obtain m<sup>6</sup>/(mol.kg<sub>cat</sub>.s).

**Table 9**  
Final values of kinetics constants vs. those found previously.

T (°C)	k <sub>1</sub>	k <sub>2</sub>	k <sub>3</sub>	k <sub>4</sub>	k <sub>5</sub>	k <sub>6</sub>
150	0.145 (0.16)	39.57 (0.78)	0.523 (0.62)	5.32 (25.4)	5.465 (4.0)	0 (6.0)
180	1.23 (2.9)	60.650 (9.5)	5.914 (3.9)	250.50 (48.4)	17.07 (93.6)	20.98 (45.2)
205	6.95 (4.8)	282.8 (66.)	13.99 (10.0)	243.30 (50.0)	19.47 (792.)	17.54 (287.)

Values in parentheses are those obtained by Osmieri et al. [5].



**Fig. 5.** Arrhenius plot for kinetic constants determined in 1 this work (L<sup>2</sup>/mol kg<sub>cat</sub> min).

It should be recalled that the mixture coefficient  $a$  is temperature-dependent; while  $b$  is not. In most cases, but not always [40],  $k_{ij}$  change slowly with  $T$ . For the PR EoS used here, we corrected  $k_{ij}$  for  $T$  when necessary (see Eq. 22). Once the mixture's  $a$  and  $b$  are calculated and the critical constants and acentric factors are known, the compressibility factor  $Z$ , necessary in the reactor model, is found by solving the cubic equation:

$$Z^3 - (1 - B)Z^2 + (A - 3B^2 - 2B)Z - (AB - B^2 - B^3) = 0 \quad (20)$$

Where,  $Z$ ,  $A$  and  $B$ , are:

$$Z = \frac{Pv}{RT} \quad A = \frac{aP}{R^2T^2} \quad B = \frac{bP}{RT} \quad (21)$$

A numerical method for solving Eqs. (20), (21), is available [42]. However, in the MATLAB<sup>®</sup> code used here, the cubic equation in  $Z$  is solved with the Cardano (trigonometric) method [41]. The method gives 3 roots. For a gaseous mixture, the largest root is chosen [22]. At this point, two additional points deserve comment.

The first comment refers to the reactor model. To calculate the component molar concentrations at a certain location in the reactor, a value for  $Z$  is required, see Eq. (5). This is calculated by solving the Eqs. (10)–(21), for the current values of  $y_i$ . The pure fluid critical constants necessary for  $a_i$  and  $b_i$ , were taken from the literature or estimated (see Table 3).  $T_c$ ,  $P_c$  and  $\omega$  are known for methanol, triolein, diolein, monoolein, glycerol, methyl oleate, and carbon dioxide. The second comment is that, for certain component pairs of the mixture the BIPs were available in the literature, but for temperatures that seldom

coincided with the reactor temperature. In these cases, we used the following regression expression, available in the process simulator, to reduce the  $k_{ij}$  to temperature  $T$ :

$$k_{ij} = k_{ij}^1 + k_{ij}^2 T + \frac{k_{ij}^3}{T} \quad (22)$$

The superscripted coefficients  $k_{ij}^n$  for use in Eq. (22), have been obtained for all the cases, so that the resulting  $k_{ij}$  can be calculated at any temperature. The use of Eq. (22) gave excellent results in our work. Other expressions are available [44].

### 3.5.1. Number of parameters and optimisation strategy

The number of BIPs for our reacting mixture has a total number of  $n = 7$  components interacting in pairs. Thus, in principle there would be a total number of 49 binary interaction parameters to be used in Eq. (16). The restrictions imposed on the values of  $k_{ij}$ , given in Eq. (16) imply that the number of different parameters is restricted to  $n(n-1)/2 = 21$  elements. Therefore there are only 21 different elements in the  $7 \times 7$  BIP matrix. This is quite a large number of parameters to calculate in an optimisation problem, unless some shortcut or specific condition can be invoked. The simplifying conditions we used consisted in taking the BIP values from phase-equilibrium studies available in the literature, and update them to the temperature of interest.

The BIPs found in the literature correspond to the following pairs: methanol-triolein, methanol-glycerol, methanol-CO<sub>2</sub>, triolein-CO<sub>2</sub>, methyl oleate-CO<sub>2</sub>, methyl oleate-methanol, and glycerol-CO<sub>2</sub>. Thus, for our problem,  $m = 7$  parameters are known in advance and the optimisation problem involves only  $n-m = 14$  unknown parameters. Additionally, there are 6 more rate parameters for each temperature. For these parameters, however, previous data have been already published, that were used as guess values in the optimisation. In summary, the strategy to solve the optimisation problem is shown on Fig. 4.

This strategy considers the joint optimisation of the PREoS parameters,  $k_{ij}$ , and the rate constants,  $k_i$ . For a given reaction temperature,  $T$ , our algorithm works as follows (refer to Fig. 4). The initial values for the  $k_{ij}$  are set to zero, while the values for the kinetic constants,  $k_i$ , are taken from our previous publications or given values with curves adjusted manually. In the first step of the algorithm (3<sup>rd</sup> block from top of Fig. 4), the  $k_{ij}$  values are those from the first step (that is, all  $k_{ij} = 0$ ) while the set of kinetic constants is now found by optimisation. Then, the  $m$  values of  $k_{ij}$  available from the literature, updated to temperature  $T$ , are used, and the global optimisation program is run to find the  $n-m$  unknown parameters with the kinetic coefficients kept constant.

The procedure is re-iterated until a satisfactory change in the value of the objective function, FO, is obtained. The values of the objective function after running the optimisation of  $k_i$ , and after running the optimisation of  $k_{ij}$ , are called FO<sub>0</sub> and FO<sub>1</sub>, respectively in Fig. 4 (see FO<sub>0</sub> and FO<sub>1</sub> in Fig. 4).

The objective function used is based on the squared relative molar concentration deviations [45]. This is:

$$FO = \sum_{i=1}^7 \sum_{j=1}^4 \left( \frac{C_{ij}^{\text{exp}} - C_{ij}^{\text{mod el}}}{C_{i,\text{max}}^{\text{exp}}} \right)^2 \quad (23)$$

Where,  $i$  is the component number, and  $j$  refers to runs at different flow rate. In this equation the values of  $C_{i,\text{max}}$  in the denominators are introduced to equalise the contributions of the minor species (triolein,

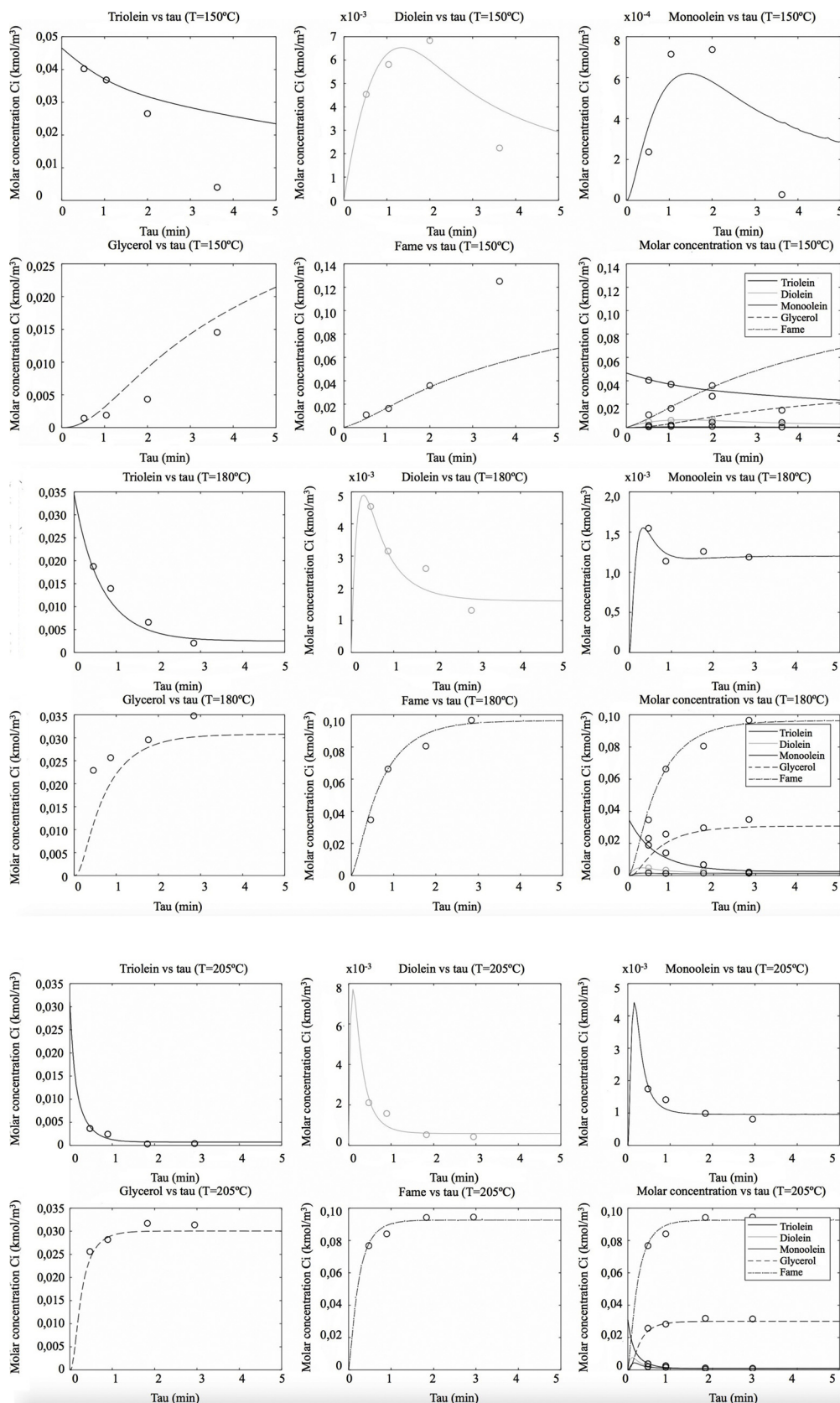


Fig. 6. Simulated and experimental concentration profiles measured at reactor exit vs. space-time,  $\tau$ . Runs at  $T = 150^\circ, 180^\circ$  and  $205^\circ\text{C}$ ,  $P = 250$  bar.

**Table 10**  
Pre-exponential factors and activation energies for kinetic parameters.

	$k_0$ $L^2/(\text{mol} \cdot \text{kg}_{\text{cat}} \cdot \text{min})^a$	E kJ/mol
$k_1$	$5.31 \times 10^{13}$ ( $9.04 \times 10^6$ )	118.07
$k_2$	$9.83 \times 10^7$ (1.5496)	52.39
$k_3$	$2.47 \times 10^{12}$ ( $4.199 \times 10^4$ )	102.13
$k_4$	$9.56 \times 10^{15}$ ( $1.625 \times 10^8$ )	121.94
$k_5$	$5.49 \times 10^5$ ( $9.33 \times 10^{-3}$ )	40.12

<sup>a</sup> Values in parentheses in  $\text{m}^6/\text{mol} \cdot \text{kg} \cdot \text{s}$ .

**Table 11**  
Pre-exponential factor and activation energy of  $k_1$  as calculated from initial rates and from optimisation.

Method	$k_0$ , $\text{m}^6/(\text{mol} \cdot \text{kg}_{\text{cat}} \cdot \text{s})$	E, kJ/mol
Initial rates	$4.76 \times 10^3$	98.51
Optimisation	$9.03 \times 10^5$	118.07

diolein, monoolein) and those of the major species (methanol and  $\text{CO}_2$ ) in the summations.

In summary: for a given temperature, 4 experimental data sets are available with 7 balance equations to be satisfied for each set, that leads to 28 equations, while 6 kinetic parameters and 14 binary interaction parameters must be determined. This means a total number of 20 parameters. Since 84 equations must be satisfied to determine 60 parameters, the use of an optimisation method such as the one described above makes sense. We want to point out that the BIPs are not constant with temperature; therefore different values for  $k_{ij}$  are determined for every temperature. As described later, when the values of  $k_{ij}$  are taken from the literature, they are updated to the temperature of interest. When the results on the  $k_{ij}$  are discussed, only those for 205 °C are given, for the reasons commented in the Results and Discussion section.

### 3.5.2. Optimisation algorithms employed

If we seek a global minimum, the alternative methods available in the MATLAB library are: *Globalsearch*, *Multistart* and *Genetic algorithm (GA)*. Many researchers agree today that GA is a successful method because it is very efficient and it supports any type of constraints. Its main drawback is the very large number of iterations. For the same biodiesel process (without catalyst), some authors [45,46] showed that the global GA and the local Levenberg-Marquardt algorithm performed quite similarly for fitting the rate constants for the batch reactor case. In the present work, we used the *Globalsearch* function because of its lower computing cost than GA. It has a reasonable speed, and it yields similarly low values for the objective function minima [47], discussed next.

Nevertheless, we wanted to investigate how GA performed in comparison. So we implemented GA and compared its results with the *Globalsearch* results for solving the parameters. We found that the two methods yielded about the same minima for FO, as shown in Table 5. Therefore, we continued with *Globalsearch* hereinafter.

**Table 12**  
Matrix of BIPs for the components of the reacting mixture containing  $\text{CO}_2$  ( $T = 205$  °C).

	Methanol	Triolein	Diolein	Monoolein	Glycerol	FAME	$\text{CO}_2$
<b>Methanol</b>	0	<b>-0.1180</b>	-0.1751	-0.1759	<b>0.0371</b>	<b>0.0485</b>	<b>0.2768</b>
<b>Triolein</b>	<b>-0.1180</b>	0	-0.1859	-0.1912	0.1456	-0.1864	<b>0.0985</b>
<b>Diolein</b>	-0.1751	-0.1859	0	-0.1865	-0.1751	-0.0862	-0.1894
<b>Monoolein</b>	-0.1759	-0.1912	-0.1865	0	-0.1901	-0.1931	-0.0513
<b>Glycerol</b>	<b>0.0371</b>	0.1456	-0.1751	-0.1901	0	-0.1886	<b>0.3093</b>
<b>FAME</b>	<b>0.0485</b>	-0.1864	-0.0862	-0.1931	-0.1886	0	<b>0.0109</b>
<b><math>\text{CO}_2</math></b>	<b>0.2768</b>	<b>0.0985</b>	-0.1894	-0.0513	<b>0.3093</b>	<b>0.0109</b>	0

Boldface  $k_{ij}$  are from the literature [50–52], and reduced to 205 °C.

## 4. Results and discussion

Table 6 and Table 7 summarise the initial mass flow rates of methanol, triolein, intermediates, and final products present at reactor outlet. We want to note here that Tables 6 and 7 supersede Table 5 of our previous paper [11], that is now amended after checking the mass balance calculations.

We proceed now to discuss the results for the kinetic parameters and for the BIPs in separate sections. Special emphasis has been put on the plausibility of the parameter values obtained.

### 4.1. Results for kinetic parameters

Table 8 summarizes our results for the kinetic constants for the 3 temperatures of the study. The rate constants are expressed in the same units as those obtained before [5] for comparison (see Table 9). The confidence intervals are also given in Table 8. This shown, the precision for rate constants is only  $\pm 30\%$ , usually larger.

The values given in Table 8 must be interpreted correctly. First, the rate constants  $k_i$  must be positive. Secondly, a zero, or close to zero, value for a reverse reaction rate constant, (for example,  $k_2$ ,  $k_4$  or  $k_6$ ) means that some reactions given by Eq. (4) tend to be irreversible.

The most relevant feature from the kinetic constants obtained is that they are similar in magnitude to those obtained previously [5], a fact that lends confidence to the values calculated in this work. Table 9 shows a comparison of the central values of Table 8 with those previously obtained [5]. In general, the central values found for the direct reaction rate constants ( $k_1$ ,  $k_3$ ,  $k_5$ ) are, in many cases, of the same order of magnitude as those previously published. Table 8 reports the confidence intervals for all the constants. It is evident that, for the 3 temperatures, the direct rate constants ( $k_1$ ,  $k_3$ ,  $k_5$ ) are relatively well determined, whereas the inverse rate constants are not so well defined. This is because the reverse rate constants ( $k_2$ ,  $k_4$ ,  $k_6$ ) had a smaller effect on the objective function used in the optimisation program. In the optimisation procedure used to obtain the central values for  $k_i$ , the restriction of positive values for the constants was indeed introduced, as noted above. However, that restriction could not be implemented in the calculation of the confidence intervals, because of the structure of the Matlab function used. Obviously, the kinetic constants must be positive. Therefore, the intervals of confidence of Table 8 are given *only to reflect* the relative precision for the constants.

As will be discussed later, the Arrhenius expressions for the kinetic constants of Table 8, and particularly their activation energies, provide sufficient basis to assure the plausibility of the values of the kinetic constant obtained. This is because only the central values were considered. Some of the reverse rate constants, for example  $k_6$ , are zero (see Table 8) for specific temperatures. This is just because the reactions written in Eq. (2) tend to be irreversible, as we will reason in the next paragraph. The tendency of  $k_4$ ,  $k_5$ ,  $k_6$  in going from 180 °C to 205 °C, anticipates smaller activation energies for these constants. In general, the confidence intervals are very large for the reverse constants.

From the values of Table 8, the equilibrium constants  $K_1$ ,  $K_2$ ,  $K_3$ , for the reactions of Eq. (2), can be figured out ( $K$  being = direct rate



constant/ reverse rate constant). For example at 205 °C, the series of values  $K_1$ ,  $K_2$ ,  $K_3$  is 0.03, 0.06, 1.0 as can be seen from the central values. In our previous publication [5] the series of values  $K_1$ ,  $K_2$ ,  $K_3$  are 0.07; 0.07; 3.0, hence, with good agreement with the values obtained here, and showing the same tendency towards irreversibility. This same feature is observed for the other reverse constants for the other temperatures. So it seems that the consecutive reactions of Eq. (2) will progressively turn from reversible to irreversible. The same feature is observed from the experimental points of  $C_A$  vs.  $\tau$ , to be presented later on Fig. 6, where triolein concentrations vs.  $\tau$  are represented. From the plots of  $C_A$  at large  $\tau$ , the equilibrium conversion of triglyceride was found to increase from 91% (at 180 °C) to 97% (at 205 °C). This would indicate that the conversion increased slightly with increasing temperature. This is the expected feature for a reversible endothermic reaction. However, here we must be very cautious when drawing conclusions because of the experimental errors associated with the measured concentrations, particularly at the lowest temperature of 150 °C.

In summary, based on the kinetic constants fitted, the three reaction steps, written as in Eq. (2), are still a little reversible. However, as the ester interchange progresses along the 3 carbons of the triglyceride molecule, the process becomes more irreversible. As temperature is increased, the equilibrium conversion would seem to increase, but more data would be needed on that. In conclusion, the values of  $k_i$  at the three temperatures allow calculation of their activation energies. The results are given in Table 9, and the regression lines are represented in the Arrhenius plots of Fig. 5.

#### 4.1.1. Activation energies of kinetic constants

The activation energies for  $k_1$ ,  $k_3$ , and  $k_4$  show higher values than the others. Moreover,  $k_2$  and  $k_5$  show a lower value compared to our previous results. However, the energies obtained are typical activation energies for catalytic processes according to the literature. A range of 60–120 kJ/mol has been given for solid-catalysed reactions by Smith [48]. Catalysis scientists [34] consider activation energies generally within a narrower range, say, about 60–80 kJ/mol, that is, at the lower end of the Smith range. For high-tonnage catalytic processes (steam-reforming, water-gas shift reaction), activation energies are in the upper end of the range (110–125 kJ/mol for industrial catalysts).

As noted above, lower activation energies of  $k_2$  and  $k_5$ , were anticipated from their values at 180 °C and 205 °C on Table 7. In summary, the activation energies obtained are typical of solid catalysts.

A possible explanation for the larger energies of Table 10 and Fig. 5, is that the uncatalysed reactions could contribute to the observed rates. Still another explanation for the low energies of Table 10, is that decomposition of FAME to free fatty acid might take place for the high temperatures of the process. In such case, less FAME than predicted by kinetics would appear, because some of the product would be lost. In a SC process with a supported catalyst, Kolb and co-workers [3] observed a substantial formation of free fatty acid at the high temperature of their process. In our experiments, free fatty acid was not examined, but our previous data indicate that this is not formed to an appreciable amount [10].

#### 4.1.2. How good are kinetic constants

We now ask ourselves how good are the estimates of the kinetic constants obtained with our non-linear optimisation method. As far as the constant  $k_1$  is concerned, however, a clear response can be given, because a linear estimate can be obtained as follows. Initial reaction rate of triolein provides an independent estimate of  $k_1$  using the differential method of kinetic analysis [38,48]. This value is written as  $k_{fit}$  in Fig. 5. Repeating this procedure for the other temperatures, the activation energy of  $k_1$  is obtained. At  $\tau = 0$ , no product is present, therefore the values of  $k_1$  can be calculated from the initial rate of reaction of A, as follows:

$$r_{\tau \rightarrow 0} \equiv r_0 = k_1 C_{A0} C_{B0}$$

$$k_1 = \frac{r_0}{C_{A0} C_{B0}} \quad (24)$$

The initial rate,  $r_0$ , can be measured and the mole concentrations  $C_{A0}$  and  $C_{B0}$ , can be evaluated exactly from the inlet compositions. A value for  $r_0$  is obtained from the initial slope of the integrated plots of  $y_A$  or  $C_A$  vs.  $\tau$ , from integral reactor data, as follows:

$$r_0 = \frac{1}{M_0} \left( \frac{dy_A}{d\tau} \right)_{\tau=0} \quad (25)$$

The values of  $k_1$  from the non-linear optimisation and those from the initial rates differ less than 18–28%, indicating that the optimisation method is acceptable.

Considering the three temperatures allows calculation of the Arrhenius equation for  $k_{fit}$ . The Arrhenius parameter values obtained with the methods are compared in Table 11 and shown in Fig. 5.

#### 4.2. Results on binary interaction parameters

The values of the Peng-Robinson interaction parameters  $k_{ij}$  are given as a matrix in Table 12, where the boldface values are the parameters taken from the literature (reduced to actual temperature). All the  $k_{ij}$  values are within the usual range for these parameters. Note however, that some parameters are negative. These correspond to the following binary systems:

- Methanol with triolein, diolein, and monoolein
- Triolein, diolein and monoolein with all the other components except CO<sub>2</sub>.
- FAME with monoolein, diolein, and triolein.
- Methanol with CO<sub>2</sub>.

A negative  $k_{ij}$  value, indicates a strong interaction between the  $i$  and  $j$  molecules in the mixture. Negative  $k_{ij}$  are usually found in mixtures of unlike molecules, i.e., polar and nonpolar molecules, showing non-ideality. In a previous article [49], we found negative BIPs in the thermodynamics of solubility of solids in neat and SC CO<sub>2</sub> modified with cosolvents. As it is known, this corresponds to cases of solubility of solids in a nonpolar gas. Obviously, we still do not have means to judge this interaction for mono- and di-olein with certain components, but we do have information about some  $k_{ij}$  for pairs involving triolein, that would be similar with the type of interaction found. For temperatures ranging from 200 °C–230 °C, Romanielo and co-workers [50] report also negative  $k_{ij}$  values for the interaction between triolein and methanol.

For 180 °C up to 205 °C, the methanol-triolein binary system has negative values for all the three  $k_{ij}$ ; with a value of  $k_{ij} = -0.1180$  found at 205 °C. It is reasonable to think that diolein and monoolein in the presence of methanol will behave similarly as triolein with methanol, in view of their similar molecular structure. This is also the case of the polar nature vs. the non-polar nature of oleins (diolein and monoolein) when mixed with glycerol (a highly polar polyol), giving negative  $k_{ij}$ . The reasoning presented above would provide some basis for the negative  $k_{ij}$  of Table 12, in particular for the pairs of the first column (or first row) and for polar/non-polar systems. Similarly, the values of the FAME-glycerides, and CO<sub>2</sub> with either diolein or monoolein should also be acceptable.

We want to make a final remark about the limit bounds imposed to  $k_{ij}$  values in the optimisation programs. Although the bounds were taken rather wide, in some cases the fitted parameters took values just onto one of the bounds, suggesting that larger (or smaller) values could be better. This behaviour is more frequent for the temperatures of 150 °C and 180 °C, as the higher experimental error or lack of accuracy of data leads the optimal parameters to fall on the bounds. Since experimental data from 205 °C fits much better to the model, best results

have been found for this temperature, while the values for the other temperatures are not so good. In general, the methodology followed in this work to obtain binary interaction parameters has proved to be very sensitive to the accuracy and error of data, so a word of caution is necessary when using the present method. The  $k_{ij}$  for 205 °C are well within expected limits, as can be noticed in Table 12.

Finally, represented on Fig. 6 are the integrated reactor model equations and the corresponding experimental data points for component concentrations, plotted vs. space-time,  $\tau$ . The values shown correspond to concentrations prevailing at reactor exit.

## 5. Conclusions

Using the experimental data for the operation of a bench-scale, integral, catalytic packed-bed reactor, the analysis and simulation of a SC biodiesel production process was done. To this end, a reactor simulation model was developed to simulate a biodiesel production process via triglyceride transesterification with methanol in the presence of CO<sub>2</sub> as cosolvent.

Previous studies show that transesterification takes place as three reversible reactions in series. For a high-pressure process, reactor simulation at a given pressure and temperature involves the use of the rate coefficients as well as the thermodynamic parameters for the fluid phase. To this purpose, the Peng-Robinson EoS, with the one-fluid Van der Waals rules for mixtures, was used. The global optimization procedure used in this work allows to simultaneously calculate the kinetic parameters,  $k_i$ , and the PR-EoS  $k_{ij}$  parameters. Based on the values of the rate constants, and on the chemical equilibrium constants calculated from them, we can conclude that the three reaction steps are still somewhat reversible, despite the excess of methanol used. As temperature is increased from 180 °C to 205 °C, the equilibrium conversion of triglyceride increases from 91 to 98%. This would suggest a reversible endothermic reaction. However, because of the experimental errors in the measured concentrations, further insight into the thermal character of the reaction is needed.

The method presented in this work allowed calculating the values of the BIPs for the intermediate reaction species (diolein, monolein) that were not available so far. The PR-EoS  $k_{ij}$  values for the component pairs obtained are reasonable on the basis of their chemical nature, but further research is needed for the full verification of the methodology. The values of  $k_{ij}$  obtained are better for 205 °C than for the lower temperatures (150 °C and 180 °C), because the method proposed in this work is very sensitive to the accuracy of the experimental data.

## Acknowledgments

The financial support of the Spanish Ministry of Economy and Competitiveness (MINECO, Madrid, project number CTQ2013-44143-R) is acknowledged. We specially thank Ms Àngela de la Rosa for help with editorial work. We are also grateful to Hecrau Group Informatic Services, S.L. (Barcelona) for using their computer facilities during part of this work.

## References

- Demirbas, Biodiesel fuels from vegetable oils via catalytic and non-catalytic supercritical alcohol transesterifications and other methods: a survey, *Energy Convers. Manage* 44 (2003) 2093–2109.
- Galia, O. Scialdone, E. Tortorici, Transesterification of rapeseed oil over acid resins promoted by supercritical carbon dioxide, *J. Supercrit. Fluids* 56 (2011) 186–193.
- J. Schürer, D. Bersch, S. Schlicker, R. Thiele, O. Wiborg, A. Ziogas, R. Zapf, G. Kolb, Operation of a small-scale demonstration plant for biodiesel synthesis under supercritical conditions, *Chem. Eng. Technol.* 39 (2016) 2151–2163.
- A. Valverde, L. Osmieri, F. Recasens, Catalytic reactor model for simulation of supercritical biodiesel synthesis with excess solvent, 10th Int. Cong. Chem. Eng. Book of Abstracts (2017) 537.
- L. Osmieri, R.A.M. Esfahani, F. Recasens, Continuous biodiesel production in supercritical two-step process: phase equilibrium and process design, *J. Supercrit. Fluids* 124 (2017) 57–71.
- A.C. Dimian, C.S. Bildea, *Chemical Process Design*, Wiley-VCH Verlag, Weinheim, 2008.
- S. Saka, Y. Isayama, A new process for catalyst-free production of biodiesel using supercritical methyl acetate, *Fuel* 88 (2009) 1307–1313.
- P. Campanelli, M. Banchemo, L. Manna, Synthesis of biodiesel from edible, non-edible, and waste cooking oil via supercritical methyl acetate transesterification, *Fuel* 89 (2010) 3675–3682.
- S. Glisic, D. Skala, The problems in design and detailed analyses of energy consumption for biodiesel synthesis at supercritical conditions, *J. Supercrit. Fluids* 49 (2009) 293–301.
- S. Glisic, D. Skala, Phase transition at subcritical and supercritical conditions of triglyceride methanolysis, *J. Supercrit. Fluids* 54 (2010) 71–80.
- J. Maçaira, A. Santana, F. Recasens, M.A. Larrayoz, Biodiesel production using supercritical methanol/carbon dioxide mixtures in a continuous reactor, *Fuel* 90 (2011) 2280–2288.
- Nafion® Sigma-Aldrich Catalog, <http://www.sigmaaldrich.com>, 2015 (Accessed 14 May 2018).
- J. Maçaira, A. Santana, A. Costa, E. Ramirez, M.A. Larrayoz, Process intensification using cosolvent under supercritical conditions applied to the design of biodiesel production, *Ind. Eng. Chem. Res.* 53 (2014) 3985–3995.
- G. Brunner, *Gas extraction, An Introduction to Fundamentals of Supercritical Fluids and the Application to Separation Processes*, Springer Verlag, Darmstadt-New York, 1994.
- M. McHugh, V. Krukonic, *Supercritical Fluid Extraction. Principles and Practice*, second ed., Butterworths, Boston, 1996.
- S.I. Sandler, *Chemical, Biochemical, and Engineering Thermodynamics*, fourth ed., John Wiley & Sons, New York, 2006.
- J.M. Smith, H.C. Van Ness, M.M. Abbott, *Introduction to Chemical Engineering Thermodynamics*, fifth ed., McGraw-Hill, New York, 1996.
- J.-N. Jaubert, L. Coniglio, The group contribution concept: a useful tool to correlate binary systems and to predict the phase behavior of multicomponent systems involving supercritical CO<sub>2</sub> and fatty acids, *Ind. Eng. Chem. Res.* 38 (1999) 5011–5018.
- D. Peng, D.B. Robinson, A new two constant equation of state, *Ind. Eng. Chem. Fund* 15 (1976) 59–64.
- O. Redlich, J.N.S. Kwong, On the thermodynamics of solutions, *Chem. Rev.* 44 (1949) 233–244.
- R.C. Reid, *Supercritical-fluid extraction. A perspective*, Hougen Lecture Series, University of Wisconsin, 1981.
- J.M. Prausnitz, R.N. Lichtenthaler, E. Gomes de Acevedo, *Molecular Thermodynamics of Fluid Phase Equilibria*, third edition, Prentice-Hall, New Jersey, 1999.
- O. Pföhl, S. Petkhov, G. Brunner, A Powerful Tool to Correlate Phase Equilibria, Herbert Utz Verlag, München, 2000.
- S.E.K. Fateen, M.M. Khalil, A.O. Elnabawy, Semi-empirical correlation for binary interaction parameters of the Peng–Robinson equation of state with the van der Waals mixing rules for the prediction of high-pressure vapor–liquid equilibrium, *J. Adv. Res.* 4 (2013) 137–145.
- J.N. Jaubert, R. Privat, F. Mutelet, Predicting the phase equilibria of synthetic petroleum fluids with the PPR78 approach, *AIChE J.* 56 (2010) 3225–3235.
- X. Xu, J.-N. Jaubert, R. Privat, P. Duchet-Suchaux, F. Braña-Mulero, Predicting binary-interaction parameters of cubic equations of state for petroleum fluids containing pseudo-components, *Ind. Eng. Chem. Res.* 54 (2015) 2816–2824.
- X. Xu, J.-N. Jaubert, R. Privat, P. Arpentier, Prediction of thermodynamic properties of alkane-containing mixtures with the E-PPR78 model, *Ind. Eng. Chem. Res.* 56 (2017) 8143–8157.
- H.P. Gros, E.A. Bottini, E. Brignole, Group contribution equation of state for associating mixtures, *Fluid Phase Equilib* 116 (1996) 537–544.
- E. Brignole, S. Pereda, Phase equilibrium engineering, in: E. Kiran (Ed.), *Elsevier SCF Science and Engineering Series*, Vol. 3 Elsevier, Amsterdam, 2013.
- A. Bertucco, P. Canu, L. Devetta, A.G. Zwahlen, Catalytic hydrogenation in supercritical CO<sub>2</sub>: kinetic measurements in a gradientless internal-recycle reactor, *Ind. Eng. Chem. Res.* 36 (1997) 2626–2633.
- E. Santacesaria, M. Di Serio, R. Tesser, M. Cozzolino, Prediction of VLE for the Kinetic Study of High-pressure Processes Based on Synthesis Gas, DGMK/SCI Conference "Synthesis Gas Chemistry" DGMK Tagungsbericht 2006-4, DGMK, Dresden, 2006.
- L. Constantinou, R. Gani, New group contribution method for estimating properties of pure components, *AIChE J.* 40 (1994) 1697–1710.
- X. Liu, R. Lin, Y. Nan, L.L. Tavlarides, Production of biodiesel from microalgae oil (*Chlorella protothecoides*) by non-catalytic transesterification. Evaluation of reaction kinetic models and phase behavior, *J. Supercrit. Fluids* 99 (2015) 38–50.
- R.J. Farrauto, C.H. Bartholomew, *Fundamentals of Industrial Catalytic Processes*, Chapman and Hall, London, 1997.
- C.-S. Tan, D.-C. Liou, Axial dispersion of supercritical carbon dioxide in packed beds, *Ind. Eng. Chem. Res.* 28 (1989) 1246–1250.
- K. Abaroudi, F. Trabelsi, B. Calloud-Gabriel, F. Recasens, Mass transport enhancement in modified supercritical fluid, *Ind. Eng. Chem. Res.* 38 (1999) 3505–3518.
- O. Levenspiel, K.B. Bischoff, Patterns of flow in chemical process vessels, in: T.B. Drew, J.W. Hoopes, Jr.T. Vermeulen (Eds.), *Advances in Chemical Engineering*, Vol. 4 Academic Press, New York, 1963.
- G.F. Froment, K.B. Bischoff, J. De Wilde, *Chemical Reactor Analysis and Design*, third ed., John Wiley & Sons, New York, 2010.
- A. Kapil, K. Wilson, A.F. Lee, J. Sadhukhan, Kinetic modeling studies of heterogeneous catalyzed biodiesel synthesis reaction, *Ind. Eng. Chem. Res.* 50 (2011)

- 4818–4830.
- [40] J.-N. Jaubert, F. Mutelet, VLE predictions with the Peng-Robinson equation of state and temperature dependent  $k_{ij}$  calculated through a group contribution method, *Fluid Phase Equilib.* 224 (2004) 285–304.
- [41] W.H. Beyer, *CRC Standard Mathematical Tables*, twenty seventh ed., CRC Press Inc., Boca Raton, (Florida), 1984.
- [42] F. Recasens, E. Velo, M.A. Larrayoz, J. Puiggené, Endothermic character of toluene adsorption from supercritical carbon dioxide at low coverage, *Fluid Phase Equilib.* 90 (1993) 265–287.
- [44] J.-W. Qian, R. Privat, J.-N. Jaubert, Predicting the phase equilibria, critical phenomena, and mixing enthalpies of binary aqueous systems containing alkanes, cycloalkanes, aromatics, alkenes, and gases ( $N_2, H_2S, H_2$ ) with PPR78 equation of state, *Ind. Eng. Chem. Res.* 52 (2013) 16457–16490.
- [45] A.M. Almargrbi, T. Hatami, S. Glisic, A.M. Orlovic, Determination of kinetic parameters for complex transesterification reaction by standard optimisation methods, *Hem. Ind* 68 (2014) 149–159.
- [46] S. Glisic, O. Montoya, A. Orlovic, D. Skala, Vapor–liquid equilibria of triglycerides–methanol mixtures and their influence on the biodiesel synthesis under supercritical conditions of methanol, *J. Serb. Chem. Soc.* 72 (2003) 13–27.
- [47] MATLAB® *Global Optimization Toolbox User’s Guide*, The Math Works Inc., 2016 (accessed 14 May 2016), [www.mathworks.com/matlabcentral](http://www.mathworks.com/matlabcentral).
- [48] J.M. Smith, *Chemical Engineering Kinetics*, third ed., McGraw-Hill, New York, 1981.
- [49] K. Abaroudi, F. Trabelsi, F. Recasens, Screening of cosolvents for a supercritical fluid. A fully predictive approach, *AIChE J.* 48 (2002) 551–560.
- [50] S. Arvelos, L.L. Rade, E.O. Watanabe, C.E. Hori, L.L. Romanielo, Evaluation of different contribution methods over the performance of Peng–Robinson and CPA equation of state in the correlation of VLE of triglycerides, fatty esters and glycerol +  $CO_2$  and alcohols, *Fluid Phase Equilib.* 362 (2014) 136–146.
- [51] W. Weber, S. Petkov, G. Brunner, Vapour-liquid-equilibria and calculations using the Redlich–Kwong-Aspen-equation of state for tristearin, tripalmitin, and triolein in  $CO_2$  and propane, *Fluid Phase Equilib.* 695 (2008) 158–160.
- [52] Z.K. Lopez-Castillo, S.N.V.K. Aki, M.A. Stadtherr, J.F. Brennecke, Enhanced solubility of hydrogen in  $CO_2$ -expanded liquids, *Ind. Eng. Chem. Res.* 47 (2008) 570–576.



### **3.2. Extraction of solid lanolin from raw wool with near-critical ethanol modified CO<sub>2</sub> —A mass transfer model**

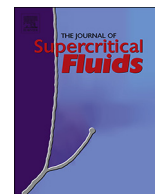
(Published in The Journal of Supercritical Fluids, 2019, 145, 151-161)

Previous works had proved the benefits of liquid or supercritical CO<sub>2</sub> solvent (in some cases a modified cosolvent) for lanolin extraction from raw wool [65] [63] [64] [66], a valuable wax with applications in pharmacy and cosmetics. Eychenne et al. [66] and patent by Bayona et al. [147] experimental data show both the extracted fraction curve at different operation conditions and the existence of a two different solubility lanolin phases. Taking this data as a background, this second paper was aimed at the development and solution of a novel transfer-mass model that correctly describes those phenomena. However, operation conditions reported presented two clear cases of study as far as modelling is concerned: a solid – near-critical extraction at 30°C and a liquid – supercritical extraction at 60°C and 80°C [66]. The complexity to model the latter case was significantly higher, as inner diffusion phenomena between different solubility lanolin phases were challenging in order to define a radial mass-transfer profile. Hence, in this paper a mass-transfer model is proposed only for the solid – near-critical extraction of lanolin from raw wool at 30°C, as lanolin layer is assumed to be a non-porous solid ensemble of two different solubility layers disposed one over the other [66]. The model was based on the shrinking-core concept described by Octave Levenspiel in his work *Chemical Reaction Engineering* [105]. To overcome the diverse solubility inside the lanolin layer, the fluid-phase concentration of lanolin in the solid-fluid interphase was modeled with a logistic function [67] in terms of those different solubilities and their respective radial boundaries.

The second milestone of the paper above is the model resolution through orthogonal collocation method. Despite not being a novel method with references such as Villadsen et al. [100] and Finlayson [102], because its faster and more accurate convergence than most finite difference methods [102] [103], it was a good chance to explore its potential in solving PDE systems as the raised ones. A rigorous comparison between Jacobi, Hermite, Laguerre, Tchebycheff and Legendre polynomials (being those latter two special cases of Jacobi polynomials) was carried out to obtain the best results (see Annex II). Finally, a modified version of even Legendre polynomials proposed by Villadsen et al. [100] and later by Finlayson [102], applied to 20 collocation points distributed throughout all the extractor length, namely the roots of the raised orthogonal polynomials [100] [101]. For this purpose, a change of variables was done to express the PDE system in terms of dimensionless variables that meet the orthogonal polynomials domain.

After a global optimization by means of Genetic Algorithm, results were found for the three parameters on which model depended, namely the solubility of each lanolin layer and the fluid phase mass-transfer coefficient. The latter agreed with those reported by Tan et al. [148] for similar systems, and solubilities met those reported by Eychenne et al. [66], showing consistency with the influence of changes in the operating conditions.

Here, although novel transfer-rate parameters for lanolin extraction in near-critical ethanol-modified CO<sub>2</sub> are given, the solving method based on orthogonal collocation highlights as an interesting contribution to non-porous solid-fluid extraction modelling. As the previous work, it was also published in *The Journal of Supercritical Fluids* (2019) by Abel Valverde and Francesc Recasens, 1<sup>st</sup> quartile in Chemical Engineering area and 2<sup>nd</sup> quartile in Physical Chemistry area, IF (2018) 3,481, with no citations so far [104].



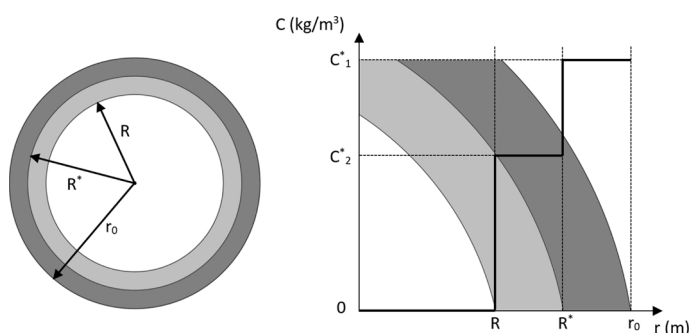
## Extraction of solid lanoline from raw wool with near-critical ethanol-modified CO<sub>2</sub> —A mass transfer model

Abel Valverde\*, Francesc Recasens

Department of Chemical Engineering, ETSEIB, UPC, Diagonal 647, E-08028, Barcelona, Spain



### GRAPHICAL ABSTRACT



### ARTICLE INFO

#### Keywords:

Solid-to-fluid  
Mass-transfer  
Near-critical CO<sub>2</sub>  
Lanoline solubility

### ABSTRACT

This study deals with the modeling of the extraction of solid lanoline from raw wool under near-critical conditions using 5% ethanol in CO<sub>2</sub>, using our previous experimental data. A mass-transfer model is developed to explain the extraction results at  $T = 30\text{ }^{\circ}\text{C}$ , below the melting point of lanoline ( $36\text{--}42\text{ }^{\circ}\text{C}$ ). Two variables are studied, the extraction pressure and the solvent mass flowrate. Our model depends on three parameters: the solubilities of the two lanoline fractions and the lanoline mass transfer coefficient. The model is a set of first-order partial differential equations, that is solved by orthogonal collocation in combination of optimization of the parameters. The fluid-side mass-transfer coefficient decreases with extraction pressure and is about  $5 \times 10^{-6}\text{ m/s}$  for  $\text{Re} < 1$  (at 70–150 bar) and depends on fluid velocity. The solubilities of lanoline fractions, independent of flowrate, agree very well with those previously reported.

### 1. Introduction

Lanoline is a natural wax secreted by the sebaceous glands of sheep that finds valuable applications in pharmacy and cosmetics [1,2]. Lanoline is traditionally obtained from raw wool by washing the sheep with water and soap. For pharmaceutical and cosmetic applications, however, it cannot be used directly, as it requires a complex purification process. An interesting alternative is to use a waterless washing

with high pressure carbon dioxide, either neat [3,4] or modified with co-solvents [5,6], in the liquid state or as supercritical fluid. This paper focuses on the high pressure extraction of lanoline from raw wool under near-critical conditions with ethanol-modified CO<sub>2</sub>, using our previously published experimental results [6,7].

Among the first authors to study SCF applied to lanoline extraction from raw wool are King and coworkers [3], Koo et al. [5] as well as New Zealand researchers [3]. These authors used neat CO<sub>2</sub> in their studies.

\* Corresponding author.

E-mail address: [abel.valverde@upc.edu](mailto:abel.valverde@upc.edu) (A. Valverde).

<https://doi.org/10.1016/j.supflu.2018.12.002>

Received 19 September 2018; Received in revised form 19 November 2018; Accepted 3 December 2018

Available online 07 December 2018

0896-8446/ © 2018 Elsevier B.V. All rights reserved.

**Nomenclature**

a	Specific surface area, m <sup>2</sup>
a <sub>i</sub> , b <sub>i</sub>	Coeffs. of polynomial P, eq. 26
C	Lanoline conc. in solvent, kg/m <sup>3</sup>
C*	Solubility, kg/m <sup>3</sup>
C* <sub>1</sub>	Solubility of fraction 1, kg/m <sup>3</sup>
C* <sub>2</sub>	Solubility of fraction 2, kg/m <sup>3</sup>
D <sub>z</sub>	Axial dispersion coeff., m <sup>2</sup> /s
k <sub>g</sub>	Fluid-side mass-transfer coeff., m/s
L	Fiber length, m
m	Total mass, kg
mf	Lanoline mass on fiber, kg
N <sub>f</sub>	Total number of fibers,
P	Pressure, bar
P <sub>n</sub> , P <sub>m</sub>	Polynomials of x
R	Final radius, m
r	Fiber radius at time t and position z, m
R*	Change radius, m
r <sub>0</sub>	Initial radius, m
R <sub>B</sub>	Bed radius, m
r <sub>v</sub>	Volumetric extraction rate, kg/m <sup>3</sup> /s
T	Temperature, °C
t	Time, s
u	Velocity, m/s
u(t)	Unit step function, eq. 25
V <sub>B</sub>	Bed volume, m <sup>3</sup>
w(x)	Weighting function, eq. 27
w'	Mass flowrate, kg/h

x	Dimensionless bed length, = z/Z <sub>t</sub>
X	Fraction extracted,
Y	Relative concentration, = C/C* <sub>1</sub>
Y*	Relative solubility, = C*/C* <sub>1</sub>
Y* <sub>12</sub>	Dimensionless variable, eq. 26
z	Coordinate from bed inlet, m
Z <sub>t</sub>	Total bed length, m

**Geek symbols**

α	Value in u(t), eq. 25
ε	Bed porosity at time t
ε <sub>0</sub>	Initial bed porosity, t = 0
η	Dimensionless radius, = r/R
η <sub>0</sub>	Initial dimensionless radius, = r <sub>0</sub> /R
θ	Solvent residence time, = Z <sub>t</sub> /u, s
ρ <sub>1</sub>	Lanoline density, kg/m <sup>3</sup>
ρ <sub>w</sub>	Wool density, kg/m <sup>3</sup>
τ	Dimensionless time
ξ	Volume. Solubility fraction = C <sub>1</sub> */ρ <sub>1</sub>
ψ	Mass transfer ratio, = k <sub>g</sub> θ / r <sub>0</sub>
δ <sub>nm</sub>	Kronecker's delta

**Acronyms**

OC	Orthogonal collocation
SCF	Supercritical fluid
SCM	Shrinking-core model

Shortly later, we filed a patent [7] based on the fractionation of lanoline lipids using compressed CO<sub>2</sub>, using either raw wool or technical lanoline as a starting material, with a solvent based on modified CO<sub>2</sub> [7]. Solvent-modified CO<sub>2</sub> under quasi-critical conditions is faster than neat CO<sub>2</sub> [6,7] requiring less pressure and temperature.

In the present paper, we specifically focus on the two lanoline fractions obtained in the process of Bayona et al. [7] and the evidence about their existence provided in our previous work [6]. Bayona found that there are two fractions: 1) A lanoline soluble in the cosolvent, that is highly polar and has a narrow molecular weight, and 2) Another lanoline fraction insoluble in the cosolvent, that has a wide molecular weight distribution. More recently, further characterization have been made [8–10].

Particularly important are the experimental results published by our group [6], as they form the basis of the present study. We carried out experiments to find the best cosolvent concentration and the extraction kinetics using subcritical and supercritical solvent. The effects of temperature, pressure, fluid velocity and wool compression were examined experimentally. In all cases, an extracted lanoline fraction appeared at low contact time, followed by a second fraction. These fractions are those found in the process of Bayona et al [7]. According to our findings [6], the best extraction conditions of raw wool are with a near-critical solvent with 5% ethanol, and with a contact time corresponding to a specific mass flowrate of 10–25 kg/h/L. Runs carried out at a very slow fluid velocity were also done by Eychenne et al. [6]. These runs in which saturation conditions prevailed, indicated the existence of the two lanoline fractions. Their previously reported solubilities [6] are one of the objectives to verify in the present work.

We had two purposes in this article. First, focusing on the two fractions of lanoline discussed above, we used the shrinking-core concept to describe lanoline extraction kinetics from raw wool. So that the solid lanoline is in contact with the fluid. Therefore, the only mass transfer resistance that exists is the individual fluid-side mass-transfer coefficient. Secondly, in order to solve the extraction model, the

orthogonal collocation method was applied to solve a system of non-linear first-order partial differential equations. This is used to determine the extracted fraction of lanoline as a function of time.

**2. Experimental background**

The experimental study behind this work was published earlier in this journal [6]. Here we report only the most significant features. The detailed process flow diagram for the high-pressure extraction setup is shown in Fig. 1. A summary of the experimental scope of the measurements, the extractor unit, the wool stock used, and the solvent properties, is given on Tables 1–3.

In our previous work [6] the effect of various variables on the extracted lanoline vs. time with modified CO<sub>2</sub> is shown. The effect of ethanol concentration, was studied (from 5% to 30% ethanol), and the

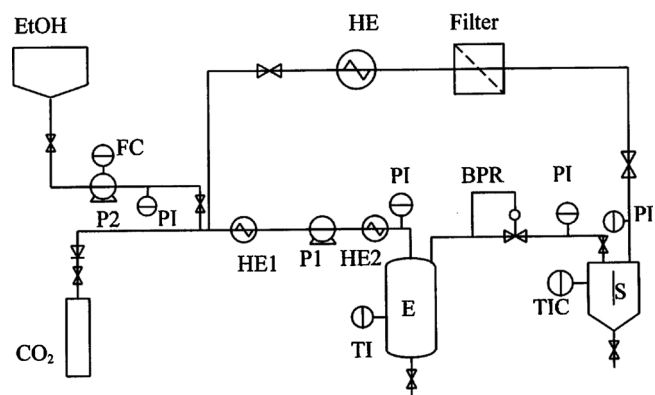


Fig. 1. Separex 200 unit process flow diagram. P1, Milton-Royal CO<sub>2</sub> pump; P2, Pulsafeeder ethanol pump; HE1, HE2, HE3, heat exchangers; BPR back-pressure control valve; E, Separex 200 extractor; S, Separex cyclone separator; PI, pressure indicators; TIC temperature indicating controller; FC, flow controller.



**Table 1**  
Scope of the measurements.

Temperature	30 °C
Pressure	70, 120,150 bar
Type of solute	Lanoline
Solvent mass flow rate, $\dot{m}$	3 – 4 kg/h
Total extraction time	8000 s
Solvent passed	0 – 6,6 kg
Solvent composition, % wt	95% CO <sub>2</sub> – 5% ethanol
Wool packing density, $\rho_B$	126,8 kg/m <sup>3</sup>

**Table 2**  
Properties of extractor and wool.

Extractor vessel	
Type	Separex SCF 200
Material	Stainless steel AISI 316L
Cylinder inside dimensions	145 mm × 30 mm (H x D)
Volume	100 cm <sup>3</sup>
Cross section	706,8 mm <sup>2</sup>
Wool fibers	
Wool load	13 g
Wool composition	60-65% wool proper, 10-15% wax and proteins, 10% soluble stains (salts), 1-20% soil and vegetable matter
Fiber geometry	Cylindrical
Fiber length (approx.)	15 cm
Diameter of fiber, D	20 μm
Lanoline content	15-30% wt
Types of lanoline	Two lanoline fractions (external and internal)
Lanoline melting point, T <sub>m</sub>	38 – 44 °C
Calculated wool fiber properties	
Initial fiber radius, r <sub>0</sub>	10 μm
Change of lanoline fraction radius, R <sup>a</sup>	9,5 μm
Final fiber radius, R	8,9 μm
Raw wool density, $\rho_w$	1314 kg/m <sup>3</sup>
Lanoline density, $\rho_l$	940 kg/m <sup>3</sup>
Bulk density, $\rho_B$	126,8 kg/m <sup>3</sup>
Bed porosity, $\epsilon$	0,9
Fiber Sauter particle size, $\delta^b$	3 × 10 <sup>-2</sup> mm

<sup>a</sup> See paragraph 3.1.

<sup>b</sup> Valverde [11].

**Table 3**  
Properties of solvent, 5%EtOH-95%CO<sub>2</sub>.

Tc	37,4 °C
Pc	77,3 bar
Solvent density, $\rho_s$	783,3 kg/m <sup>3</sup>
Viscosity, $\mu^a$	0,0648 mPa.s

<sup>a</sup> Fields et al. [12].

lower % of ethanol chosen for economy. The raw wool employed contained 15% wt of lanoline. In our previous paper, it was reasoned that extraction of lanoline takes place in two successive steps. This corresponds to the two lanoline fractions located one over the other on the wool fibers, rather than to a wax mixture or fractions adsorbed on different sites. This is consistent with the fractions observed earlier in our patent [7] and in [6], as reasoned in the Introduction section.

In a previous paper [6], we showed that the best temperature for desorption of lanoline is 30 °C. At this temperature, the lanoline layer on wool is a solid and the fluid is a near-critical liquid. On Fig. 2 is the PT diagram for the 5% ethanol-95% CO<sub>2</sub>. For 30 °C and pressures of 70–150 bar, the solvent is a liquid. In the present paper, extraction at a single temperature of 30 °C is considered, where extraction is fastest, and the lanoline layer is a solid in contact with the fluid.

### 3. Extraction model

#### 3.1. Previous findings to consider in the model

As will be discussed in the model, the radius of the wax layer where lanoline fraction changes, is a key parameter of the model. Eychemen et al. [6] performed extraction runs at very low velocity runs to determine the approximate lanoline solubility at different pressure and temperature conditions. By plotting the mass of solute extracted vs. mass of fluid passed, the slope of the graph allows to estimate the approximate solubilities of lanoline fractions in the solvent. From these experiments, it can be clearly seen the two lanoline fractions, suggesting that one desorbs after the other, and that the first fraction has a larger solubility than the second one. The break point of the total curve with different slopes, can be used to calculate the radius at which the change of fraction occurs. On Fig. 3, we give the regression lines of the equilibrium plots for 30 °C and 150 bar.

Fig. 3 provides the radius of the lanoline, R\*, where the change occurs, by the following calculation. Let N<sub>f</sub> be the total number of wool fibers in the extractor, m<sub>wool</sub> = 40 g, and the lanoline mass extracted at the break point, m<sub>extracted</sub> = 2892 g (from Fig. 3). Therefore, for a cylindrical wool fiber, we can write:

$$\frac{N_f \pi (r_0^2 - R^{*2}) \rho_l}{N_f \pi r_0^2 \rho_w} = \frac{m_{\text{extracted}}}{m_{\text{wool}}} \quad (1)$$

Hence,

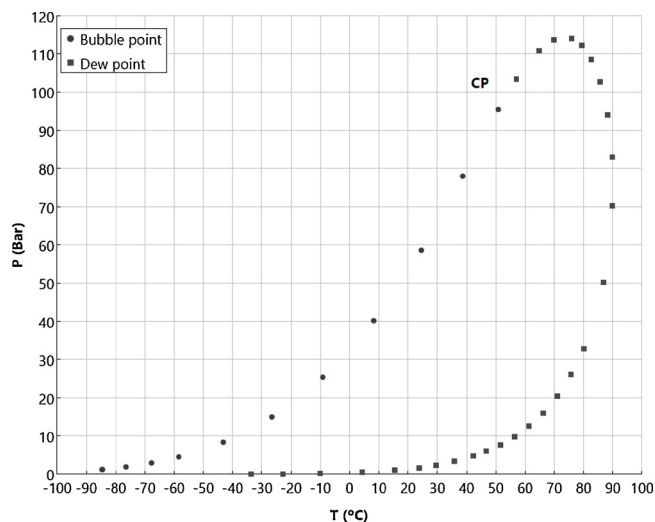
$$R^* = r_0 \sqrt{1 - \frac{m_{\text{extracted}} \rho_w}{m_{\text{wool}} \rho_l}} = 10 \sqrt{1 - \frac{2,892 \cdot 1314}{40 \cdot 940}} = 9,48 \mu\text{m}$$

Also, a simple calculation allows to establish the final radius, R, at which the lanoline has been completely extracted. The raw wool is known to contain, on average, 15–30% wt, of lanoline. After a few runs with the model it is clear that the exact % is in our case 15%, so that, m<sub>lanoline</sub>/m<sub>wool</sub> = 015. Then, we have:

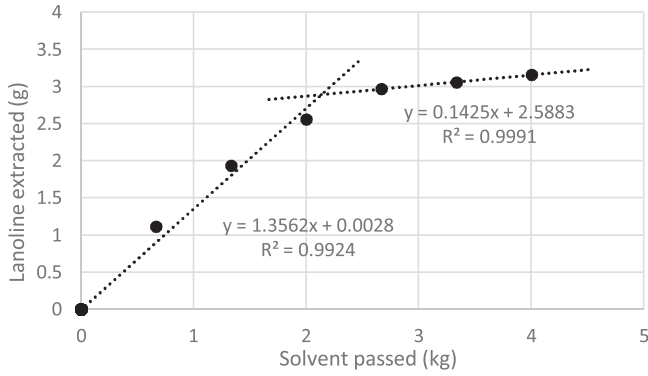
$$\frac{N_f \pi (r_0^2 - R^2) \rho_l}{N_f \pi r_0^2 \rho_w} = \frac{m_{\text{lanolin}}}{m_{\text{wool}}} \quad (2)$$

From which,

$$R = r_0 \sqrt{1 - \frac{m_{\text{lanolin}} \rho_w}{m_{\text{wool}} \rho_l}} = 10 \sqrt{1 - 0,15 \frac{1314}{940}} = 8,89 \mu\text{m}$$



**Fig. 2.** PT diagram for the mixture 5% ethanol-95% CO<sub>2</sub>. CP is the critical point. For T = 30 °C and pressures above 70 bar, the fluid is a liquid.



**Fig. 3.** Extraction run at very small fluid velocity. Data points are from [6], showing the extraction of the two lanoline fractions. Regression of solubility data to obtain the intersection point ( $T = 30\text{ }^\circ\text{C}$ ,  $P = 150\text{ bar}$ , wool packing density  $390,3\text{ kg/m}^3$ ). The slopes of the straight lines give the solubilities in g/kg fluid.

### 3.2. Mass conservation equations for lanoline

#### 3.2.1. Conservation in the fluid phase

We first write the conservation equation for lanoline solute in the fluid phase, and then, on the solid phase. Let us call  $r_v$  the volumetric rate of lanoline desorption in the fluid. For the general case of disperse fluid flow in the  $z$ -direction, the conservation of lanoline over a differential bed length  $dz$ , leads to the following equation:

$$\varepsilon \frac{\partial C}{\partial t} + u \frac{\partial C}{\partial z} = D_z \frac{\partial^2 C}{\partial z^2} + r_v \quad (3)$$

Where  $C$  is the lanoline concentration in the fluid at a position  $z$ , and time  $t$ . For a packed bed of 15 cm long, with wool of equivalent particle size of less than 1 mm (see Table 1), the equivalent number of perfect mixers in series is very large, and, consequently,  $Pe_z \gg 100$ , so plug flow conditions can be assumed. So, Eq. (3) can be written as:

$$\varepsilon \frac{\partial C}{\partial t} + u \frac{\partial C}{\partial z} = r_v \quad (4)$$

Where  $r_v$  is the solid-to-fluid mass transfer rate:

$$r_v = k_g a (C^* - C) \quad (5)$$

In which, it is assumed that the fluid-phase concentration of lanoline close to the solid in contact with the solvent is  $C^*$ , and  $a$  is the specific surface of wool fibers per unit volume of bed. If  $r$  is the cylindrical fiber radius and  $L$  its length, the area  $a$ , can be written as:

$$a = \frac{2\pi r L}{\pi r^2 L} (1 - \varepsilon) = \frac{2(1 - \varepsilon)}{r} \quad (6)$$

With  $a$  from Eq. (6) into Eq. (4), this becomes:

$$\varepsilon \frac{\partial C}{\partial t} + u \frac{\partial C}{\partial z} = k_g \frac{2(1 - \varepsilon)}{r} (C^* - C) \quad (7)$$

Where  $r$  is a function of time and position in the bed  $z$ , i.e.,  $r(t, z)$ . Note that the void fraction,  $\varepsilon$ , in Eq. (7) is in fact variable with time because the lanoline dissolves into the fluid leaving a thinner film.

#### 3.2.2. Conservation of lanoline in the wool phase

To describe extraction kinetics, we apply the concept of the shrinking-core model (SCM) with a cylindrical geometry. Refer to the radius shown on Fig. 4.

Whereby, the lanoline mass on a fiber of length  $L$ , at time  $t$  and position  $z$  in the bed, will be:

$$m_f = (\pi r^2 - \pi R^2) L \rho_l \quad (8)$$

The number of fibers initially loaded in the bed (assumed cylindrical) is given by:

$$N_f = \frac{(1 - \varepsilon_0) \pi R_B^2 Z_t \rho_w}{\pi r_0^2 L \rho_w} = \frac{(1 - \varepsilon_0) R_B^2 Z_t}{r_0^2 L} \quad (9)$$

Where  $R_B$  is the bed radius and  $Z_t$  is the bed length. In Eq. (9),  $\varepsilon_0$  is the initial bed porosity, and  $r_0$  the initial raw wool fiber radius. See other symbols in the Nomenclature section. With this, the initial mass of lanoline present in the packed bed, will be:

$$m = N_f m_f = \frac{\pi \rho_l (1 - \varepsilon_0) R_B^2 Z_t}{r_0^2} (r^2 - R^2) \quad (10)$$

Differentiating  $m$  with respect to  $t$  in Eq. (10), and dividing by the total bed volume,  $V_B$ , we will have:

$$\frac{\partial m}{V_B \partial t} = \frac{2\rho_l (1 - \varepsilon_0) r}{r_0^2} \frac{\partial r}{\partial t} \quad (11)$$

Now, equating Eqs. (11) and (5) with a minus sign, we obtain the change of radius  $r$  with time, as:

$$\frac{\partial r}{\partial t} = -\frac{k_g (1 - \varepsilon) r_0^2}{\rho_l (1 - \varepsilon_0) r^2} (C^* - C) \quad (12)$$

In summary, the model equations for the extraction of the lanoline is represented by the following system of equations:

$$\frac{\partial C}{\partial t} = \frac{2k_g (1 - \varepsilon)}{r \varepsilon} (C^* - C) - \frac{u}{\varepsilon} \frac{\partial C}{\partial z} \quad (13a)$$

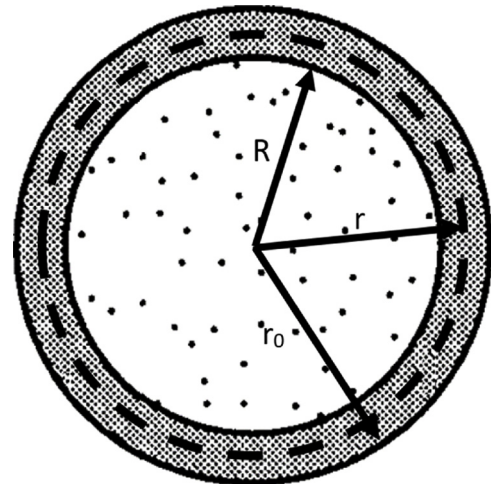
$$\frac{\partial r}{\partial t} = -\frac{k_g (1 - \varepsilon) r_0^2}{\rho_l (1 - \varepsilon_0) r^2} (C^* - C) \quad (13b)$$

A refinement of the model is possible. The value of  $\varepsilon_0$  is the initial porosity of the wool bed employed to calculate  $N_f$  with Eq. (9). Note, however, that the bed void fraction increases because of the extraction of the lanoline due to its extraction to the fluid. In the above equations,  $\varepsilon$  is the porosity at time  $t$  and radius  $r$ . The two porosities are related through the following equations:

$$\varepsilon_0 = 1 - \frac{\rho_B}{\rho_w} \quad (14)$$

$$\varepsilon = 1 - \frac{\rho_B r^2}{\rho_w r_0^2} \quad (15)$$

With these expressions substituted into Eq. (13), the model equations become:



**Fig. 4.** Shrinking-core model for a cylindrical lanoline cover on wool fiber. Initial external radius is  $r_0$ , inner radius is  $R$ , and variable dotted line is the shrinking radius  $r$ .

$$\frac{\partial C}{\partial t} = \frac{2k_g(1 - \varepsilon_0)r(C^* - C) - ur_0^2 \frac{\partial C}{\partial z}}{r_0^2 - (1 - \varepsilon_0)r^2} \tag{16}$$

$$\frac{\partial r}{\partial t} = -\frac{k_g}{\rho_l}(C^* - C) \tag{17}$$

This is a system of two coupled nonlinear partial differential equations.

For the initial conditions, it is considered that initially the extractor contains no fluid, so that the lanoline concentration in the fluid is zero at all positions, and the initial radius of lanoline is  $r_0$  everywhere. In this idealized model, when solvent is allowed into the bed ( $t = 0$ ), an instantaneous compression of the bed takes place with  $C = 0$ . For the boundary conditions, it is assumed that at bed entrance ( $z = 0$ ) the fluid has no lanoline, so  $C = 0$ . Therefore, the initial and boundary conditions for  $C(t, z)$  and  $r(t, z)$ , are:

$$\begin{aligned} C(0, z) &= 0 \\ r(0, z) &= r_0 \\ C(t, 0) &= 0 \end{aligned} \tag{18}$$

### 3.3. Dimensionless mass transfer model

In order to make the model non-dimensional we define the following variables:

$$\begin{aligned} Y &= \frac{C}{C_1} \\ Y^* &= \frac{C^*}{C_1^*} \\ \eta &= \frac{r}{R} \\ x &= \frac{z}{Z_t} \\ \tau &= \frac{t}{\theta} \end{aligned} \tag{19}$$

Where the first lanoline fraction, whose solubility is  $C^*_1$ , is the first solute to be extracted. With the new variables, the model becomes:

$$\frac{\partial Y}{\partial \tau} = \frac{2\psi(1 - \varepsilon_0)\eta(Y^* - Y) - \eta_0 \frac{\partial Y}{\partial x}}{\eta_0^2 - (1 - \varepsilon_0)\eta^2} \tag{20}$$

$$\frac{\partial \eta}{\partial \tau} = -\xi\psi(Y^* - Y) \tag{21}$$

Where  $Y$  is the solute concentration in the fluid,  $\eta$  is the fiber radius relative to  $R$ , and  $\tau$  is the dimensionless time. See the other symbols in the Nomenclature. The initial and boundary conditions for Eqs. (20) and (21), are now:

$$\begin{aligned} \text{IC:} \\ Y(0, x) &= 0 \\ \eta(0, x) &= \frac{r_0}{R} \end{aligned} \tag{22}$$

$$\begin{aligned} \text{BC:} \\ Y(\tau, 0) &= 0 \end{aligned} \tag{23}$$

The system of Eqs. (20) to (23) allows the method of orthogonal collocation to be applied in the  $x$ -domain  $[0,1]$ , thus simplifying the solution (see paragraph 4). We now describe the conditions at the fluid-solid interphase as respect to the two lanoline fractions.

### 3.4. Lanoline concentration profile and the change radius called $R^*$

As has been reasoned, the lanoline fractions is deposited on the wool fiber forming two coaxial cylindrical layers. These have different but well-defined solubilities. At  $T = 30^\circ\text{C}$ , they are solid. Neglecting the solubility of the fluid in the solid, diffusional effects in the lanolin layer do not take place. Refer now to Fig. 5.

Based on the position of solute on the fiber, a concentration profile of  $C^*$  in a double step is considered, in such a way that  $C^*_1$  is constant from the initial radius  $r_0$ , up to the change radius  $R^*$ , where the second lanoline fraction is reached. From this radius inwards, the saturation value  $C^*_2$  is constant again up to the point where the radius occupied by lanoline becomes zero. In order to mathematically include the step function in the saturation values, the following continuous function is used [13]:

$$u(t) = \frac{1}{1 + e^{-\alpha t}} \tag{24}$$

The use of the exponential function is justified by the need to introduce a step expression with a continuous derivative so numerical problems are avoided. In Eq. (24), an appropriate large value for  $\alpha$  is used. In the non-dimensional model, Eqs. (20) and (21), the core radius  $\eta$ , takes values comprised between  $r_0/R$  and 1, while the fluid concentration  $Y$ , takes values between 1 and 0. Therefore, the initial step for the extraction will take a value  $C^*_1/C_1^*$  during the first fraction extraction, hence,  $u(t) = 1$ . The second step will take a value  $C^*_2/C_1^*$ , when the radius is  $\eta = R^*/R$ .

In order to introduce a double step like that of Fig. 5 in the model equations, it is necessary to write a value  $Y^*(\eta)$  containing the two different limits. This is done in the following way:

$$Y^*(\eta) = \left( \frac{1 - Y^*_{12}}{1 + e^{-\alpha(\eta - \eta^*)}} + Y^*_{12} \right) \left( \frac{1}{1 + e^{-\alpha(\eta - 1)}} \right) \tag{25}$$

Where  $Y^*_{12} = C^*_2/C^*_1$ , and  $\eta^* = R^*/R$ . In the calculations, a large value for  $\alpha$  is given, for example,  $\alpha = 10^{1000}$ .

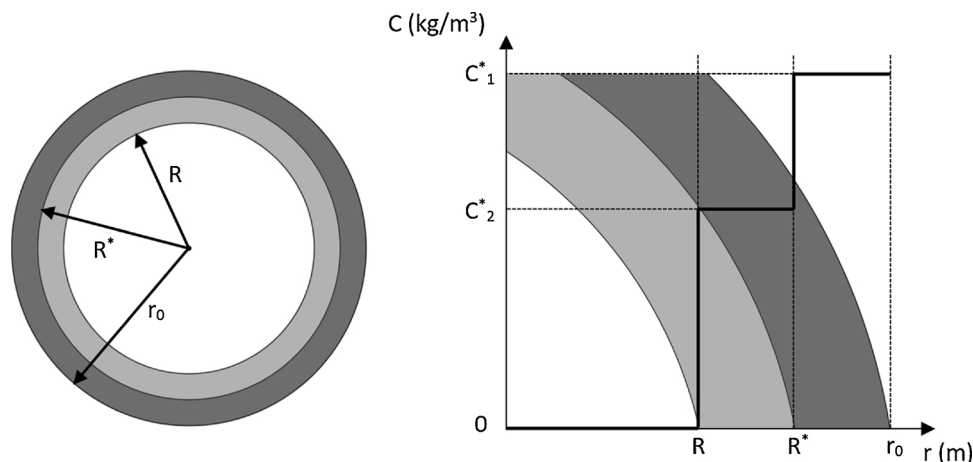


Fig. 5. Solid lanoline solubility profile on the fiber. The change radius  $R^*$ , separates the two lanoline fractions.

**Table 4**

Effect of pressure on extracted fraction vs. time. Experimental data from Eychenne et al. [6]. (T = 30 °C, flowrate  $m' = 4$  kg/h,  $\rho_B = 126,8$  kg/m<sup>3</sup>).

Solvent passed [kg]	Lanoline extracted fraction		
	70 bar	120 bar	150 bar
0,67	0,24	0,28	0,34
1,33	0,45	0,53	0,61
2,0	0,64	0,68	–
2,67	0,70	0,78	0,86
3,33	0,76	0,84	0,93
4,0	0,85	0,87	0,97

## 4. Solution of extraction model and parameter optimization

### 4.1. Orthogonal collocation

In order to solve Eqs. (20) to (23), the method of orthogonal collocation (OC) [14–16] is proposed. As shown by early authors, the OC method is faster and more accurate than most finite difference methods for PDE systems [15,16]. The method is based on approximating the trial solution consisting in a sum of orthogonal polynomials with initially unknown coefficients. We express the solution to  $Y$  and  $\eta$  as a function of time  $\tau$  only. In such a way that when the trial solutions are substituted in the system, the partial derivatives with respect to  $\tau$  become total derivatives and the partial derivatives with respect to  $x$ , which are  $Y$  and  $\eta$ , form a matrix of values for different points of the bed (the collocation points). As a result, the system of equations become a system of initial-value ODEs with respect to  $\tau$  that is readily solved with standard methods (Runge-Kutta, Gear, etc). The trial functions in our case can be written in terms of the polynomials, in a simple manner, as:

$$\begin{aligned} Y(\tau, x) &= \sum_{i=0}^N a_i(\tau) P_i(x) \\ \eta(\tau, x) &= \sum_{i=0}^N b_i(\tau) P_i(x) \end{aligned} \quad (26)$$

The orthogonality condition for two polynomials with respect to a weighting function,  $w(x)$ , in the interval  $[a, b]$ , is:

$$\int_a^b w(x) P_n(x) P_m(x) dx = C_{(n)} \delta_{nm} \quad (27)$$

Where  $C_{(n)}$  is a normalizing factor, and  $\delta_{nm}$ , is Kronecker's delta. In our case, only the polynomials derived from the Rodrigues' formula [18] were considered, so it guarantees the orthogonality condition. In practice, the Legendre polynomials, orthogonal in the  $[0,1]$   $x$ -interval, were used. Its series expression has been used here to calculate the coefficients with MATLAB (see dimensionless model in paragraph 3.3).

The form of the trial functions used in the present work are those used by Villadsen and Stewart [14] and later by Finlayson [15]. They are:

$$\begin{aligned} Y(\tau, x) &= Y(\tau, 1) + (1 - x^2) \sum_{i=0}^{N-1} a_i(\tau) P_i(x^2) \\ \eta(\tau, x) &= \eta(\tau, 1) + (1 - x^2) \sum_{i=0}^{N-1} b_i(\tau) P_i(x^2) \end{aligned} \quad (28)$$

In these expressions the polynomials are the same but in terms of  $x^2$ , so even polynomials are obtained, hence with symmetrical solutions. Therefore, they are useful in problems with a symmetric geometry (slab, cylinder, sphere). The number of collocation points is  $N + 1$ , that is, equal to the number of polynomials employed. The collocation points are the roots of the polynomial of highest degree [14,17]. In this case,  $x = 0$  and  $x = 1$  are both taken also as collocation points, as a boundary condition is set at  $x = 0$  and the solution is wanted at  $x = 1$ . In our calculations, we employed the Legendre polynomials with even exponents, because they give the least oscillations, and they are easily constructed with MATLAB.

### 4.2. Optimization of parameters

After the solution of the extraction model is obtained in terms of the fluid phase concentration at bed exit, it is necessary to characterize the system parameters from the experimental data available [6]. These results are usually given in terms of the lanoline fraction extracted as a function of time at the bed exit, hereinafter called  $X(t)$ . To calculate  $X$  from the solution of the model, the following expression is used:

$$X(t) = \frac{\int_0^t C(t, Z_i) dt}{\int_0^\infty C(t, Z_i) dt} \quad (29)$$

To calculate the fraction extracted as a function of time, the following change of scale is done:  $m_s = m' t$ , where  $m'$  is the mass flow rate of solvent.

The optimization function used to obtain the parameters, is that employed previously by Valverde et al. [19], based on the squared relative deviations of extracted fractions, as:

$$FO = \sum_{i=1}^n \left( \frac{X_i^{exp} - X_i^{model}}{X_{max}^{exp}} \right)^2 \quad (30)$$

In Eq. (30),  $n$  is the number of data points at a given pressure (for 70 and 120 bar,  $n = 6$ ; and for 150 bar,  $n = 5$ ). The global optimization algorithm used was *Genetic Algorithm* (GA) available in MATLAB, because the local optimization methods performed quite inefficiently. GA instead provides precise and optimal results with problems with poor flexibility as discussed elsewhere [19].

## 5. Results and discussion

### 5.1. Experimental data available

The experimental data used for modeling are those of Eychenne et al. [6]. Mainly, from the Figs. 4, 6a and 9 a, at 30 °C, of that article [6]. Shown on Tables 4 and 5, are the Eychenne results for extraction at different pressures and mass flowrates. In what follows, we first discuss the parameters and identify their relative sensitivity. Then, we discuss the results obtained for the fitted parameter values as a function of the extraction pressure. And finally we present the effects of flowrate on the parameters.

### 5.2. Model parameters and their sensitivities

In our system model, Eqs. (21)–(24), there are three parameters to fit together. Apart, there is the change radius  $R^*$ , that is well established (see Paragraph 3.1). The three unknown parameters are the values of the solubilities of the lanoline fractions  $C^*_{1}$  and  $C^*_{2}$ , and the value of the fluid-side mass transfer coefficient  $k_g$ . It is interesting to note that in the present problem the geometry and size of wool fibers are well defined, so that a value of  $k_g$  is obtained separately from the transfer area. Approximate values of the lanoline solubilities are available (see Fig. 3)

**Table 5**

Effect of solvent flowrate on extracted lanoline vs. Time. Experimental data from Eychenne et al. [6]. (T = 30 °C, P = 120 bar,  $\rho_B = 126,8$  kg/m<sup>3</sup>).

Solvent flowrate			
3 kg/h		4 kg/h	
Solvent passed [kg]	Lanoline extracted fraction	Solvent passed [kg]	Lanoline extracted fraction
0,75	0,46	0,67	0,32
1,5	0,59	1,33	0,55
2,25	0,77	2,0	0,71
3,0	0,8	2,67	0,81
		3,33	0,84

that can be used as a guide to verify the validity our results.

We first examine the sensitivity of the objective function FO, Eq. (30), to the parameters. The results can be seen on Fig. 6 where the response of the fraction extracted is simulated when parameters are changed one at a time.

Fig. 6 reflects their relative influence. The value of  $C^*_1$  shows a large effect on the first part of the response curve, while  $C^*_2$  shows it higher effect on the second part of the curve. This was expected from the two lanoline fraction involved. Also, it is seen that when  $k_g$  and the solubilities are both diminished, the response curve is shifted to the right or to larger extraction times. The most important feature of Fig. 6 is the individual effect of each parameter.

The order of importance is as follows. A slight change in  $C^*_2$  has the larger effect on the model solution, so this parameter is the most sensitive. The next influential parameter is  $k_g$ , and finally the least influential is  $C^*_1$ , but it is still important. As a consequence, the response curve is good enough for determining the three model parameters as all three affect the extraction rate.

### 5.3. Results on the pressure dependence

Combining the solution of the model with parameter optimization, the resulting values of  $k_g$ ,  $C^*_1$  and  $C^*_2$  are obtained, and given on Table 6, where the effect of extraction pressure is clearly seen.

As can be observed, the values of the solubilities  $C^*_1$  and  $C^*_2$  at 150 bar, agree very well with those measured previously [6]. The effect of pressure on the solubilities between 70 and 150 bar is the expected result. Also, the fluid-side individual mass transfer coefficient  $k_g$ , is seen to decrease with increasing pressure, in agreement with Brunner [20], and to those found experimentally by other authors [21,22].

It is interesting to note that  $k_g$  has been obtained independently of the transfer area. Note also that its measurement is free from other mass transfer resistances that usually accompany mass transfer studies (see liquid-to-supercritical fluid mass transfer studies, [21]). The value of  $k_g$  obtained can be compared using the publication of Puiggené et al [21] where the global  $K_g$  measured by several SCF authors are represented as a function of Reynolds number. The range of the particle Reynolds number in our experiments [6] is around  $Re \sim 0,7$  ( $Re = 065-0,72$ ) with a value of the  $k_g$  (Table 6) in the order of  $5 \times 10^{-6}$  m/s (for

**Table 6**

Final values of parameters at 70, 120 and 150 bar (30 °C and 4 kg/h flow rate).

	70 bar	120 bar	150 bar
$k_g$ (m/s)	$5,66 \cdot 10^{-6}$	$4,08 \cdot 10^{-6}$	$1,59 \cdot 10^{-6}$
$C^*_1$ (g/kg solvent)	0,665	0,807	$1,178 (1,35)^b$
$C^*_2$ (g/kg solvent)	0,194	0,242	$0,427 (0,47)^b$
FO <sup>a</sup>	0,00071	0,00315	0,00105

<sup>a</sup> Minimum of objective function. <sup>b</sup>Values in parenthesis are the solubility values measured by Eychenne et al. [6] (see Fig. 3).

$P = 70-120$  bar). This value is similar to the mass transfer coefficient found by Tan et al. [22], also at the interphase solid-fluid, for the case of packed-bed desorption with SCF.

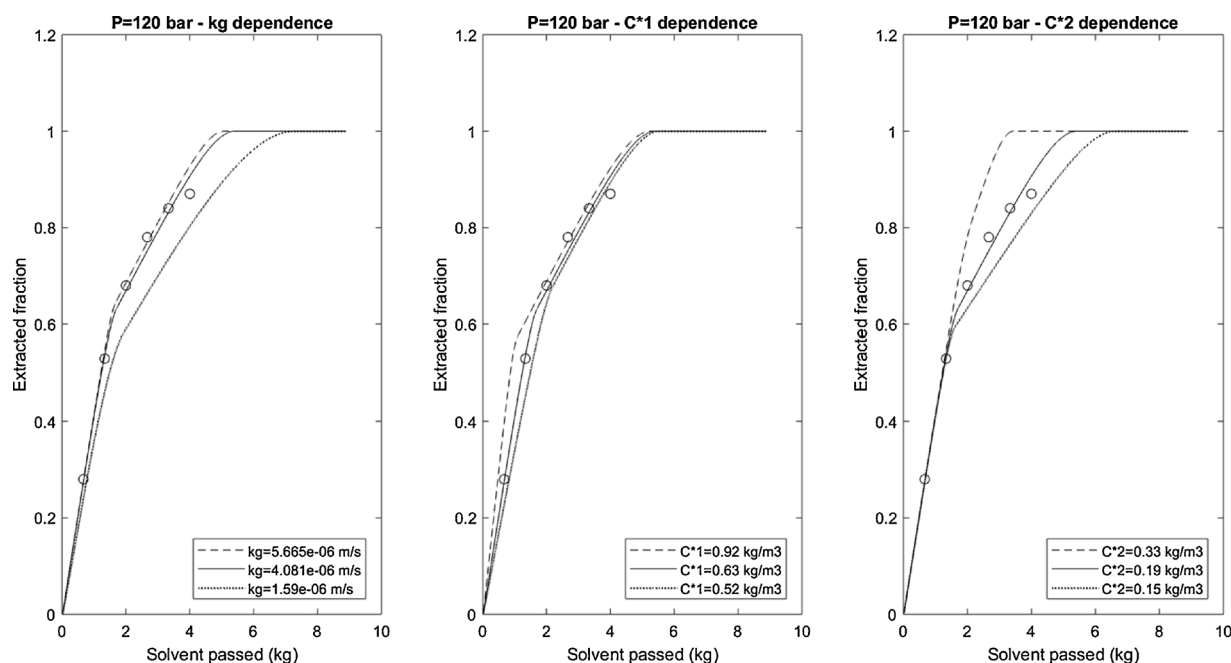
On Fig. 7, the evolution of the lanoline fluid concentration and the wool fiber radius as a function of time, are presented for the pressures of the study, using 20 different collocation points along the bed. On the left of the figure, are the graphs of  $C(t, z)$  vs. time, and on the right are the fiber radius  $r(t, z)$ , using the Legendre polynomials.

Some features of Fig. 7 deserve comment. Because the first lanoline fraction extracted is more soluble, the radius drops sharply (see left side of Fig. 7) for a change radius of  $R^* = 9,5$  mm. This sharp drop is a characteristic feature of SCM kinetics for gas-solid reactions [24]. Because the second lanoline fraction is less soluble, the radius  $r(t)$  for  $r = R^*$  drops also to a constant final value for  $r = R$  (all lanoline extracted). This is seen very clearly on Fig. 7.

As regards to the lanoline concentration in the fluid, graph of  $C$  vs.  $t$ , the stepped profile at bed exit is due to the steps in the radius and to the solubility. These steps can be smoothed by using more collocation points in the model. However, the concentration graph at bed exit, presents a response with more steps.

In Fig. 8 the steps have been eliminated, and continuous lines for the concentration and the radius vs. time can be seen. The general trend with extraction pressure is clearly apparent on Fig. 8.

As a final result, the corresponding extracted fraction of lanoline vs. solvent passed is given on Fig. 9, where the effect of pressure on extraction kinetics is calculated with our model with the optimum parameter values of Table 6, and with the experimental data [6]. As can be observed on Fig. 9, the best extraction conditions at 30 °C correspond to 150 bar.



**Fig. 6.** Simulated extracted fraction vs. solvent passed when parameters are varied ( $T = 30$  °C,  $P = 120$  bar,  $m' = 4$  kg/h). Data points are from [6].

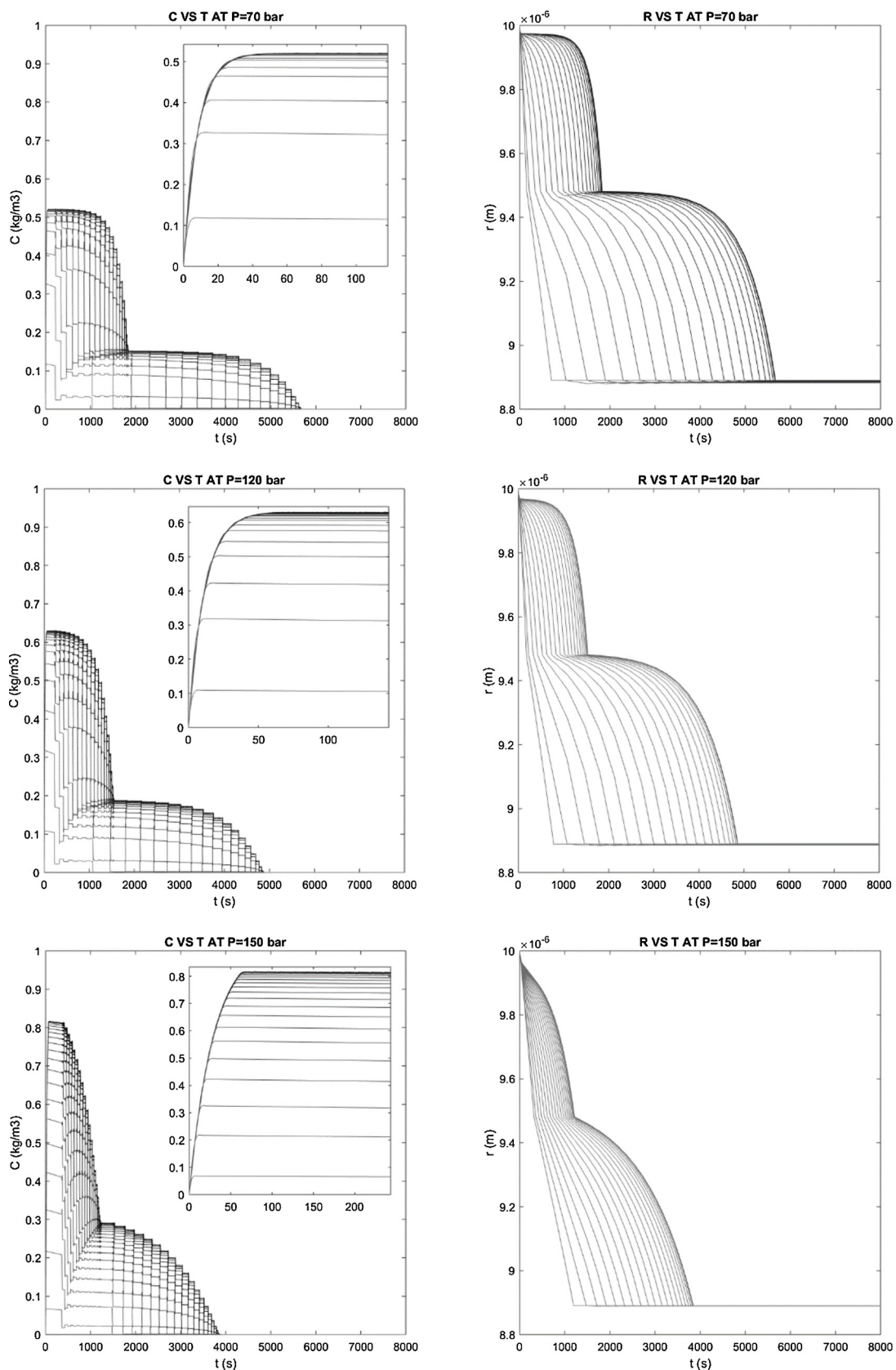


Fig. 7. Time evolution of lanoline concentration in the solvent (left) and radius of the fiber (right), for extraction at 30 °C and  $w' = 4$  kg/h flow rate, for 20 different collocation points along the bed. The leftmost lines (clearer lines) show the evolution of lanoline concentration at a point near the extractor entrance while the right lines (darker ones) are for points near the exit.

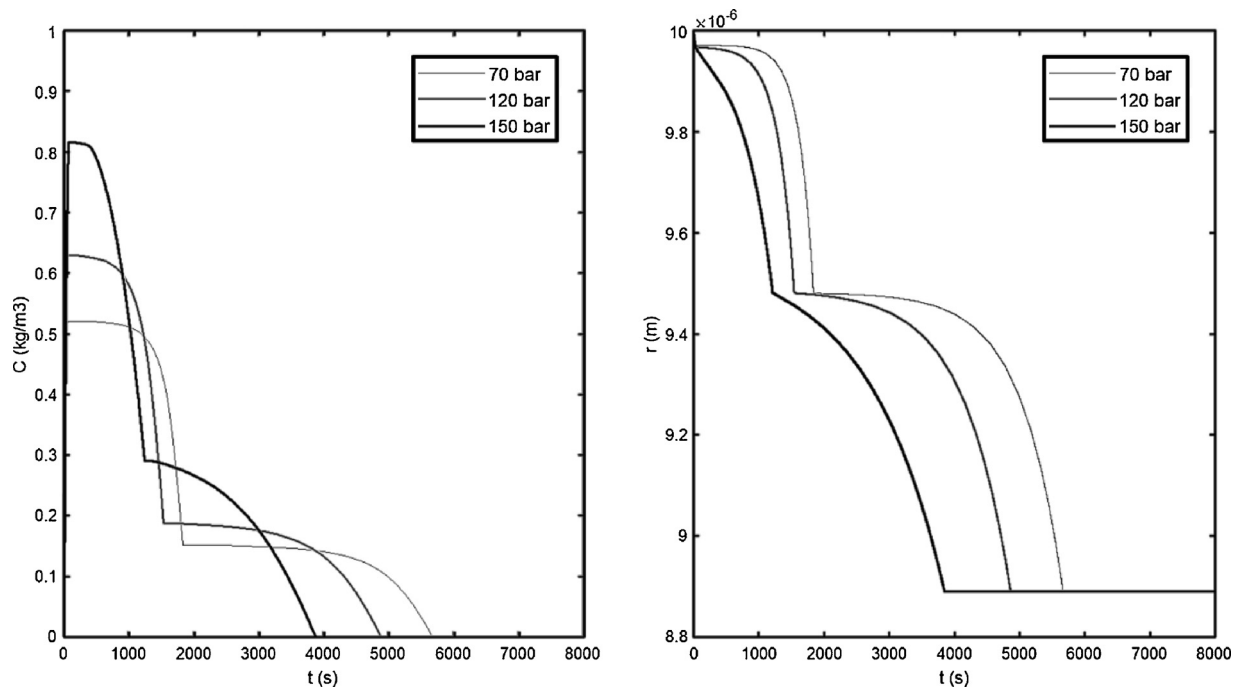


Fig. 8. Calculated concentration of lanoline in the solvent (left) and radius of the fiber (right) evolution with time at the extractor exit, as a function of pressure (30 °C and  $m' = 4$  kg/h).

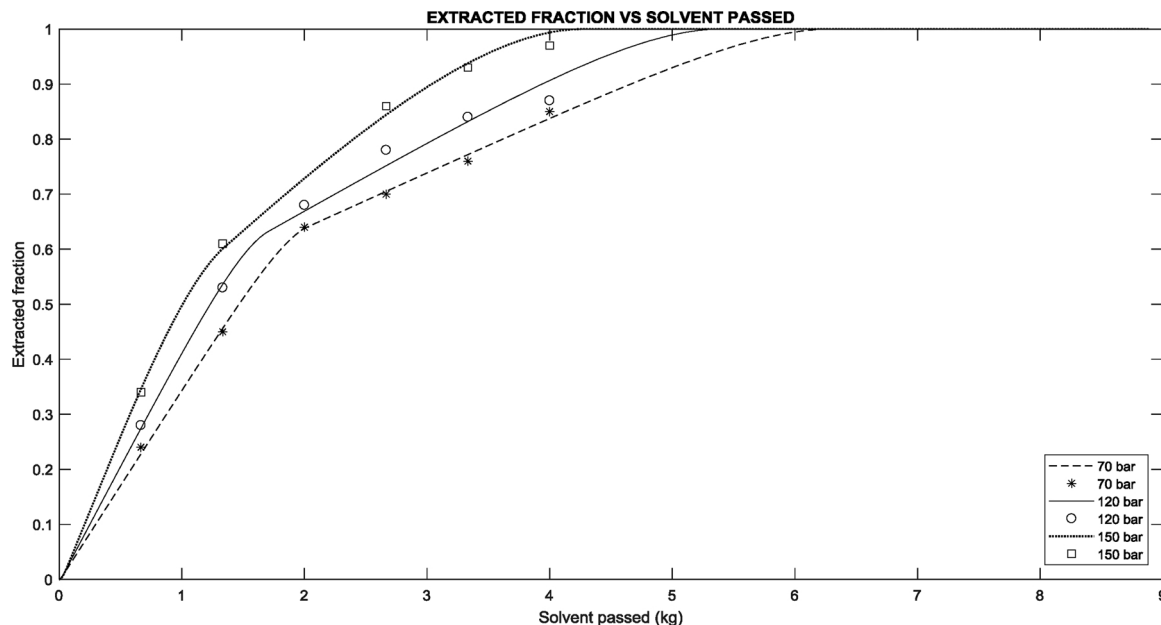


Fig. 9. Fraction extracted vs. solvent passed at 70 bar, 120 bar and 150 bar, (30 °C and  $m' = 4$  kg/h) Data points are from [6], continuous lines with our model with best parameter values (Table 7). Best extraction conditions are 30 °C and 150 bar.

Table 7

Effect of flowrate on optimized parameters for 3 kg/h and 4 kg/h flow rate (120 bar and 30 °C.).

	3 kg/h	4 kg/h
$k_g$ (m/s)	$4,09 \cdot 10^{-6}$	$5,13 \cdot 10^{-6}$
$C_1^*$ (g/kg solvent)	1,149	0,853
$C_2^*$ (g/kg solvent)	0,236	0,256
OF	0,00703	0,00422

5.4. Results on flowrate dependence

After obtaining the parameters  $k_g$ ,  $C_1^*$ , and  $C_2^*$  by modeling the experimental results of Eychenne et al [6], the fitting of the model parameters was done for the two fluid velocities for which extraction data are also available. The results are summarized on Table 7. These show the effect of increasing the solvent flowrate from 3 kg/h to 4 kg/h for runs at constant T and P (30 °C and 120 bar).

The solubility values  $C^*$  should be independent of the flowrate because solubilities are thermodynamic properties, thus they depend on temperature and pressure only. This is the behavior observed for  $C_1^*$  ( $C_1^* = 0,242$  g/kg, Tables 6 and 7), hence with good agreement. For

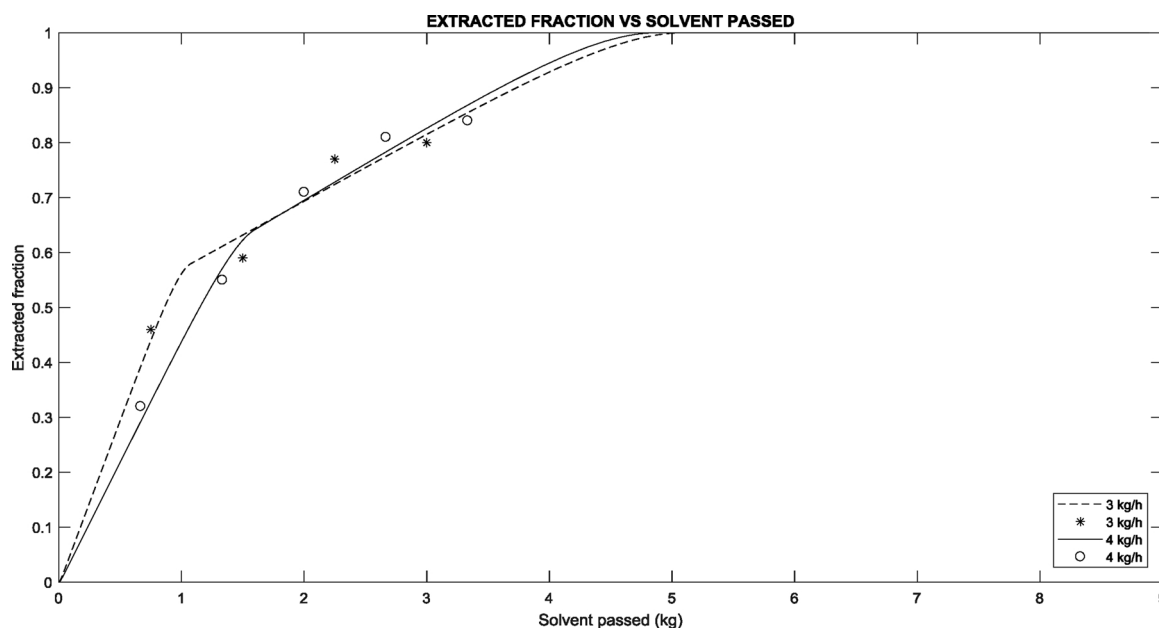


Fig. 10. Effect of flowrate of the fraction of lanoline extracted vs. solvent passed (3 and 4 kg/h, 30 °C and 120 bar). Experimental data points are from [6], continuous lines with our model with best parameters of Table 7.

the case of  $C^*_2$  the agreement is fair, as there is a slight difference between the values of Tables 6 and 7 ( $C^*_2 = 0,807$  g/kg in Table 6 vs  $C^*_2 = 0,853$ – $1,149$  g/kg in Table 7). Anyway, these values agree within an order of magnitude.

As regards to the mass transfer coefficient  $k_g$ , this is seen to increase by 25% with an increase in fluid velocity of 33%, see Table 8. The observed increase implies a dependence of  $k_g$  on velocity as  $u^{0,77}$ . In general, for gas flow and liquid flow in packed beds, the dependence of a fluid side mass transfer coefficient  $k_g$  on the velocity is taken into account through correlations of the type  $Sh$  vs.  $Re$ . For SCF, for example, Tan et al. [22] found a correlation of the type,  $Sh = 0,38 Re^{0,83} Sc^{1/3}$  ( $Re = 2$ – $40$ ) with an exponent 0,83, larger than the exponent accepted today for packed beds (0,65) but closer to our exponent 0,77. Unfortunately, we do not have enough results at larger velocities to establish the velocity exponent with more precision.

On Fig. 10, the fraction of lanoline extracted vs. time for increasing mass flowrates, is presented [6]. As reasoned above, a dependence of  $k_g$  on  $u^{0,77}$  would require to examine a wider flowrate range to see its effects reflected on the extraction rate, that are hardly seen on Fig. 10.

## 6. Conclusions

Using our previous publications, on the experimental results of near-critical extraction of lanoline from raw wool with modified carbon dioxide, a mass transfer model has been developed that explains extraction kinetics at 30 °C, below the melting point of lanoline. That is, at conditions where the lanoline is a nonporous solid and the solvent is a compressed liquid. Our model assumes that the two lanoline fractions deposited one over the other on the wool fibers. This picture is suggested by our previous findings and a patent. When extraction starts, the first lanoline fraction to be dissolved is that on the outer layer in close contact with the fluid. Afterwards, a less soluble second fraction is extracted. The first fraction has a higher solubility, while the second one is less soluble. A critical or change radius, denoted  $R^*$ , that characterizes the transition from one fraction to the other, was measured and is taken as a key data in the model.

The extraction model proposed is based on the shrinking-core model for solid-fluid reactions, with a cylindrical particle geometry. Because, the mass transfer area for lanoline dissolution is relatively well defined, the only rate coefficient for lanoline extraction is the individual fluid-

side coefficient,  $k_g$ , a value for which was calculated here. The extraction model depends on two additional parameters. These are the solubilities of the two lanoline fractions. The model performs very well in describing the extraction kinetics observed, in particular, the dependence on the extraction pressure and the solvent mass flowrate.

In order to solve the system of nonlinear first-order differential equations that appear in the model, the method of orthogonal collocation, as developed by the early authors (Villadsen and Stewart, and Finlayson), has been used here. In this way, we have chosen the Legendre polynomials that gives a solution with least oscillation. The solution of the extracted lanoline and the radius of remaining lanoline on the fiber as a function of time, allows to obtain the model parameters. These are obtained using a global optimization procedure combined with the solution of the mass transfer model.

## References

- [1] J. Thewlis, Lanoline for Cosmetic Applications, Agro food industry hi-tech, 1977, pp. 14–20 may-june vol..
- [2] G.R. Whalley, Take a closer look at lanoline, TAPPI J. 81 (5) (1998) 11–18.
- [3] F.W. Jones, D.R. Bateup, S.R. Dixon, S.R. Gray, Solubility of wool wax in supercritical carbon dioxide, J. Supercrit. Fluids 10 (1995) 100–111.
- [4] M. Cygnarowicz-Provost, J.W. King, W.N. Marmar, P. Magidman, Extraction of woolgrease with supercritical carbon dioxide, J. Am. Oil Chem. Soc. 71 (1994) 222–225.
- [5] B.S. Kuo, J.C. Kim, J.H. Jeon, H.K. Bae, Desorption of wool grease from greasy wool with supercritical carbon dioxide, Hwahak Konghak 30 (1992) 491–498.
- [6] V. Eychenne, S. Sáiz, F. Trabelsi, F. Recasens, Near-critical solvent extraction of wool with modified carbon dioxide—experimental results, J. Supercrit. Fluids 21 (2001) 23–31.
- [7] J.M. Bayona, P. Erra, Z. Moldovan, C. Domínguez, E. Jover, F. Recasens, M.A. Larrayoz, Method for obtaining lipid fractions from wool or lanoline using pressurized carbon dioxide, Patent WO 2002100990A1, Oficina Española de Patentes y Marcas, Madrid, filed 16.07.2005.
- [8] C. Domínguez, E. Jover, F. Garde, J.M. Bayona, P. Erra, Characterization of supercritical fluid extracts from raw wool by TLC-FID and GC-MS, J. Am. Chem. Oil Soc. 80 (2003) 717–724.
- [9] C. Domínguez, P. Erra, J.M. Bayona, Physico-chemical and dyeing properties of raw wool extracted by pressurized CO<sub>2</sub>/Modifiers, Text. Res. 80 (2010) 651–659.
- [10] M. López-Mesas, F. Carrillo, M.C. Gutiérrez, M. Crespi, Alternative methods for wool wax extraction from wool scouring wastes, Aceites y grasas 58 (2007) 402–407.
- [11] A. Valverde, PhD Thesis, in preparation, 2018.
- [12] P.R. Fields, T.L. Chester, A.M. Stalcup, Viscosity estimation in binary and ternary supercritical fluid mixtures containing carbon dioxide using a supercritical fluid chromatograph, J. Liquid Chrom. Rel. Technol 34 (2011) 995–1003.
- [13] M. Fullana, F. Trabelsi, F. Recasens, Use of neural net computing for statistical and



- kinetic modelling and simulation of supercritical fluid extractors, *Chem. Eng. Sci.* 54 (1999) 5845–5862.
- [14] J.V. Villadsen, W.E. Stewart, Solution of boundary-value problems by orthogonal collocation, *Chem. Eng. Sci.* 22 (1967) 1483–1501.
- [15] B.A. Finlayson, Packed bed reactor analysis by orthogonal collocation, *Chem. Eng. Sci.* 26 (1971) 1081–1091.
- [16] M. Morbidelli, A. Servida, G. Storti, R. Paludetto, S. Carrà, Application of the orthogonal collocation method to some chemical engineering problems, *Eng. Chim. Ital.* 19 (1983) 47–60.
- [17] A.S. Olagunju, L.J. Folake, M.T. Raji, Comparative study of the effect of different collocation points on Legendre-collocation methods of solving second-order boundary value problems, *IOSR J. Math.* 7 (2013) 35–41.
- [18] M.R. Spiegel, *Advanced Mathematics for Engineers and Scientists*, McGraw-Hill, New York, 1971.
- [19] A. Valverde, L. Osmieri, F. Recasens, Binary interaction parameters from reacting mixture data. Supercritical biodiesel process with CO<sub>2</sub> as cosolvent, *J. Supercrit. Fluids* 143 (2019) 107–119.
- [20] G. Brunner, *Gas Extraction*, Steinkopff Darmstadt, Springer, New York, 1994.
- [21] J. Puiggené, M.A. Larrayoz, F. Recasens, Free liquid-to-supercritical fluid mass transfer in packed beds, *Chem. Eng. Sci.* 52 (1997) 195–212.
- [22] C.-S. Tan, S.-K. Liang, D.-C. Liou, Fluid-solid mass transfer in a supercritical fluid extractor, *Chem. Eng. J.* 38 (1980) 17–23.



### **3.3. A Nonlinear Autoregressive Exogenous Neural Network Applied to Weibull Distributed Data. Lanolin Supercritical Extraction Model.**

(Submitted to SN Applied Sciences)

In the previous paper, the model for temperatures above 30°C had not been studied due to the complexity that diffusion phenomena inside the lanolin layer with diverse solubilities supposed. It is, indeed, because at 60°C and 80°C the extraction of lanolin from raw wool with modified CO<sub>2</sub> (5% ethanol) occur at supercritical conditions in the solvent phase whereas solute is a complex liquid mixture. Thus, a neural network was designed with the purpose to predict the extracted fraction of lanolin at different operation conditions and times. Data available from Eychenne et al. [66] depends on five operation variables: temperature (60-80 °C), pressure (120-200 bar), solvent mass flow rate (3-5 kg/h), wool packing density (127-318 kg/m<sup>3</sup>) and time (~1h). However, the scarceness of data presented a serious handicap for a proper performance of the neural model. This problem was overcome by means of a statistical fitting of the experimental data based on Weibull distribution, achieving this way the enough augmentation of data for the correct performance of the neural network. The election of Weibull distribution is theoretical based on the demonstrated expression of the concentration of lanolin in the solvent at the extractor's inlet. This expression is a Weibull type distribution over time, and, despite modelling the concentration behavior at the inlet of the extractor, it was assumed that the same Weibull type distribution would model the extractor outlet, as the breakthrough curve shape evolves over time and extractor's height maintaining always this sigmoid function (see Annex III – On Inlet Breakthrough Curve's Demonstration). Weibull statistical distribution demonstrated to be very suitable for extraction breakthrough curves

modelling, showing both accuracy and the usual shape of this kind of processes (see Annex III – On Weibull distribution results).

Once semi-experimental data was ready, the work was focused on finding out the most appropriate neural network architecture for the problem. As MATLAB software is suited with a sophisticated neural networks packaging, diverse MATLAB already defined architectures were applied, such as *feedforwardnet* and diverse recurrent neural networks (NARNET, NARXNET). First exploration on simple feedforward neural networks was motivated by the work of Fullana et al. [67], but those demonstrated to not be well suited for time series prediction, as its simple architecture, where response moved upwards from an initial layer to an output through diverse hidden layers, was not practical for previous time values feedback. Hence, taking Bulsari and Palosaari work as a reference [69], recurrent neural networks for time series were studied, resulting a Nonlinear Autoregressive Neural Network with External Input (NARXNET) the most suitable architecture for the aim of the work. Roughly, it consists of a feedforward architecture with an exogenous time independent input and a time dependent input formed by the output values at previous times, namely feedback inputs, with a delay vector applied to both inputs. Temperature, pressure, solvent mass flow rate and wool packing density were set as exogenous inputs. Therefore, taking first extracted fraction value as feedforward input and a unitary time delay, network could be trained to meet next extracted fraction values available using Levenberg-Marquardt (LM) algorithm [69]. The matrix input structure, and the number of hidden layers and nodes was crucial for obtaining the best convergence as it was discussed in the paper. Moreover, due the huge variability of the data, each training run converged to a different minimum (remember that LM is a local optimization algorithm), so an iterative algorithm was implemented in order to meet a

tolerance restriction in the resulting error between both the training and the testing experimental data (one extracted fraction curve at a certain operation condition was saved and not used in the training step), and the predicted values (see Annex III – On Solving Method Abstract).

Final results showed that a (5,3,1) nonlinear autoregressive neural network could predict the experimental data with a  $\pm 0.42\%$  error, and, in addition, potential optimization applications were possible with a fast convergence, as, once neural network is trained and ready to predict, its computational cost for iterative optimization algorithms is minimum. Some of those applications were proved using Genetic Algorithm optimization methods.

This paper is currently under review since 3<sup>th</sup> of April (2020) in *SN Applied Sciences*, a new Springer Nature journal with no IF available so far [62].



# Hybrid Nonlinear Autoregressive Neural Network – Weibull Statistical Model Applied to the Supercritical Extraction of Lanolin from Raw Wool

Abel Valverde\*, Jesus Alvarez-Florez, Francesc Recasens  
Chemical Engineering Department, Universitat Politècnica de Catalunya,  
Av.Diagonal 647, 08028 Barcelona, Spain

\*Corresponding author: *abel.valverde@upc.edu*

## Abstract

Supercritical extraction of lanolin from raw wool with modified CO<sub>2</sub> (5% ethanol) at temperatures above the melting point of lanolin ( $T = 36\text{-}42\text{ }^{\circ}\text{C}$ ) is difficult to model because of the multicomponent diffusion in the liquid layer. In this work, a neural network model is proposed based on the experiments previously published by our research group. Experimentally, the extraction of a 100-cm<sup>3</sup> packed bed of raw wool depends on five variables, i.e., temperature (60-80 °C), pressure (120-200 bar), solvent mass flow rate (3-5 kg/h), wool packing density (127-318 kg/m<sup>3</sup>), and time (~1h). A nonlinear autoregressive exogenous (5,3,1) neural network was designed and trained with the experimental data augmented using an empirical Weibull statistical function. This correctly predicts the lanolin breakthrough at the extractor exit with only  $\pm 0.42\%$  error. The simple arithmetic of neural network allows a fast optimization with Genetic Algorithm to find optimum operation conditions for the extraction process.

**Nomenclature:** See end of the manuscript

## 1. Introduction

The early attempts to extract lanolin from raw wool using supercritical carbon dioxide date back to 1990s, and are due to King and coworkers [1], Koo et al [2], and to New Zealand researchers [3]. In all cases, pure compressed CO<sub>2</sub> was employed. The first experimental systematic study, published a few years later, is that due to Eychenne et al. [5]. In their work, CO<sub>2</sub> modified with 5% ethanol was used under near-critical conditions, this allowed operation at lower pressure. Higher extraction yields were possible. Recently, we published an article dealing with the modelling of the extraction of solid lanolin from raw wool using modified CO<sub>2</sub> [6]. At 30 °C, the lanolin is a solid (mp. 38-44 °C) and the fluid is a liquid. At these conditions, extraction kinetics could be interpreted with the shrinking-core concept [6]. In this case, the pure solid lanolin covering the wool fibers is in direct contact with the fluid and is progressively dissolved by direct contact with the flowing solvent. In this model, two solid lanolin fractions are assumed to be located one over the other on the wool fiber, as suggested by the experimental results obtained in the fractionation process developed by Bayona et al. in a patent [6,7], and chemical characterization of the fractions [8,9].

The present work, based also on the same experiments [5], comprises wider ranges of temperature, pressure and other variables, as summarized on Table 1. As can be seen, in all the extraction runs lanolin is extracted from a liquid into a near-critical fluid at temperatures above the melting point of lanolin, i.e., 60 to 80°C. For these temperatures, the conservation equations for lanolin become complicated by the fact that a multicomponent liquid phase is present, involving additional mass transfer resistances compared with the extraction at temperatures below the lanolin melting point where lanolin is a pure solid. For temperatures



above 30 °C, the lanolin diffusion equations from the liquid up to the liquid-fluid interface makes solution very difficult.

**Table 1. Settings for lanolin extraction runs from raw wool [5]**

Temperature, T	60, 80 °C
Pressure, P	120, 150, 200 bar
Solvent mass flow rate, Q	3, 4, 5 kg/h
Solvent passed	up to 5 kg
Wool packing densities, $\rho_B$	127, 159, 227, 318 kg/m <sup>3</sup>

\* Extraction solvent composition: 95% wt. CO<sub>2</sub> – 5% ethanol; total extraction time 1 h.

In [11], an exact analytical solution of the conservation equations with some simplifying assumptions was obtained. The corresponding theory is summarized below in the Conservation equations, paragraph 3.1. It is evident, however, that the conservation equations become very complex, because not only lanolin is extracted to the fluid, but also CO<sub>2</sub> and ethanol are dissolved in the liquid phase, thus giving a very complicated situation to model.

On the other hand, artificial neural networks (ANNs) can be used to solve the conservation equations in cases where the physical model is either very complex, the solution algorithm is difficult, or does not exist.

In the present case, we are interested on the effects of the factors that affect the yield and rate of extraction of lanolin from wool. The main feature of ANNs is their ability to learn by reproducing the output from a given input by observation and minimization of the error providing an optimal set of weighting parameters. In the prediction process, no explicit law is assumed that relates output from input.

It will be apparent that our extraction runs can somehow be considered as a desorption process. As regards to a work on adsorption column operation, an early paper due to Bulsari

and Palosaari [15, 17] dealing with system identification was very interesting to us. In our case, lanolin is not held on wool fibers by adsorption. Instead, dissolution of lanolin in the solvent depends on how far the fluid is from saturation, so the rate of dissolution depends on a mass transfer coefficient and a concentration gradient.

Our objective in this paper was to use neural network computing combined with the properties of the conservation equations for lanolin extraction from raw wool, as described below in the Theory section. To this end, experimental data about the variables that affect the extraction rate and yield are essential. In this sense, the systematic extraction study data [5] is most appropriate to try to model.

## **2. Experimental background**

The experimental study behind this work was published elsewhere (Eychenne et al., 2001); the reader should refer to it for details. Here, we report only a few significant features. A Separex 200 unit was used for the high-pressure extraction. This is shown on Fig 1. A summary of the scope of the measurements, features of the extractor vessel and the wool stock used are given in Table 2.

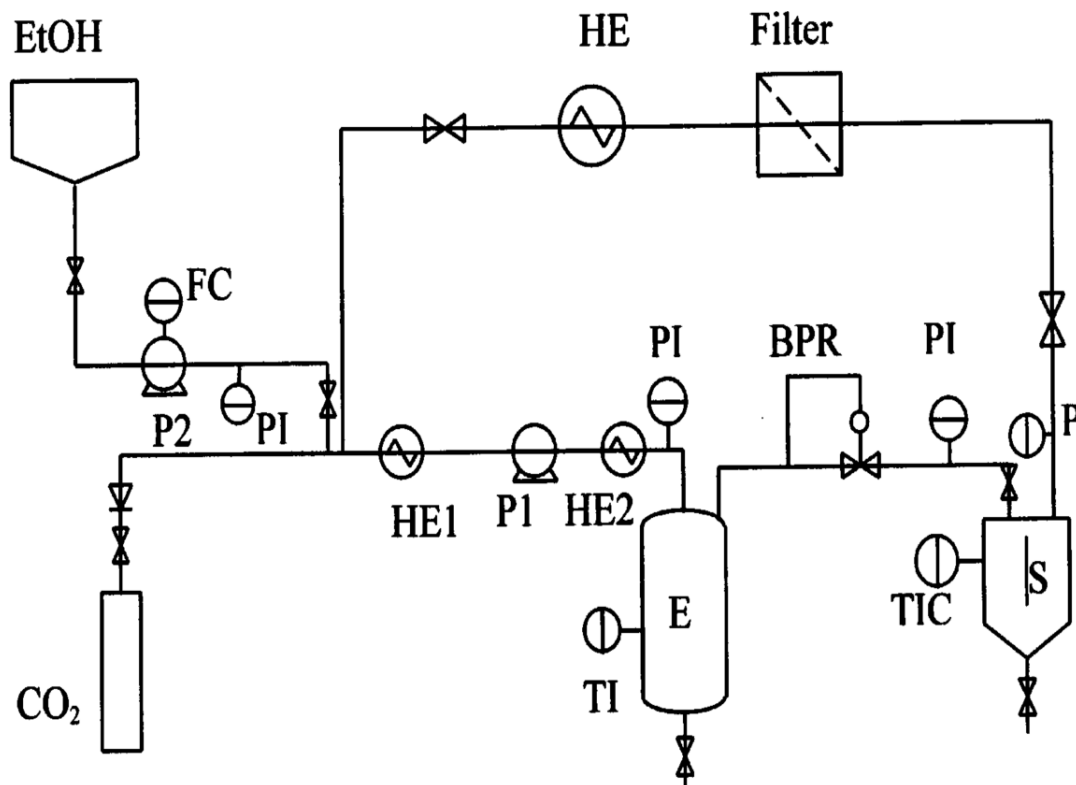


Fig 1. Separex 200 unit process flow diagram. P1, Milton-Royal CO<sub>2</sub> pump; P2, Pulsa-Feeder ethanol pump; HE1, HE2, HE3, heat exchangers; BPR back-pressure control valve; E, Separex 200 extractor vessel; S, Separex cyclone separator; PI, pressure indicators; TIC temperature indicating controller; FC, flow controller

Table 2. Properties of extractor and wool used [5]

Extractor vessel	
Shape	Cylindrical
Material	Stainless steel AISI 316L
Inside dimensions	145 mm x 30 mm (H x D)
Volume	100 cm <sup>3</sup>
Cross section	706.8 mm <sup>2</sup>
Wool fibers	
Wool load in extractor	13 g
Wool composition	60-65% wool proper, 10-15% wax and proteins, 10% soluble stains (salts), 1-20% soil and vegetable matter
Fiber geometry	Cylindrical
Average fiber length (approx.)	15 cm
Initial fiber radius, R	8.3 μm
Initial fiber radius, r <sub>0</sub>	10 μm
Lanolin content	20 % wt
Types of lanolin	Two lanolin fractions (external and internal)
Lanolin melting point	38 – 44 °C

### 3. Theory

#### 3.1 Conservation equations.

Consider a vertical bed of packed wool from which the lanolin deposited on the fibers is extracted by the flowing solvent. See item E, in Fig 1. At  $t = 0$ , a step input of lanolin-free solvent is introduced in the bed at  $z = 0$ , while fluid saturated with lanolin is removed at the other end of the bed at  $z = z_T$ . If pressure drop is neglected, the solvent passes through the bed in plug flow at constant pressure and temperature. The rate of extraction depends on two factors: the solubility of lanolin in the solvent and a mass transfer coefficient.

Two assumptions are made: 1) since the amount of lanolin on the wool is small, the bed void fraction,  $\epsilon$ , is about constant and equal to that of packed wool; and 2) since the layer of liquid lanolin covering the fibers is very thin, the total mass transfer area,  $a$ , can be considered constant; therefore, the holdup of liquid in the bed,  $\epsilon_L$ , can be taken equal to its average value during the extraction process. With these assumptions, the balance equations for lanolin in the flowing fluid and in the liquid phase, respectively, are:

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = r_v \quad (1)$$

$$-\frac{\partial(V_L C_L)}{\partial t} = V k_G a (C_g^* - C_g) \quad (2)$$

Where  $r_v$  is the volumetric mass transfer rate at position  $z$  in the bed and time  $t$ , and  $k_G$  is the solvent-side film mass transfer coefficient. In eqn. (2),  $C_g^*$  is related to the concentration of lanolin in the liquid,  $C_L$ , through a Henry-type equilibrium constant, as is usually done in high pressure studies. Then:

$$C_g^* = K C_L \quad (3)$$

Note that  $C_L$  decreases with increasing  $z$ , because of lanolin dissolution in the solvent occurs for increasing values of  $z$ . Using the following relationship:

$$V_L = V \epsilon_L \quad (4)$$

the gas and liquid holdups,  $\epsilon$  and  $\epsilon_L$ , and the Henry constant,  $K$ , eqns. (1) and (2), become:

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = K k_G a (C_L - \frac{C_g}{K}) \quad (5)$$

$$-\frac{\partial C_L}{\partial t} = \frac{K k_G a}{\epsilon_L} (C_L - \frac{C_g}{K}) \quad (6)$$

Where  $\epsilon_L$  is the liquid holdup taken as an average during a run. If the bed is initially loaded with new wool, and the inlet solvent is lanolin-free, the initial and boundary conditions for eqns. (5) and (6), are:

$$C_g(t, z = 0) = C_g(t = 0, z) = 0 \quad (7)$$

$$C_L(t = 0, z) = C_{L0} \quad (8)$$

Where  $C_{L0}$  is the initial concentration of lanolin covering the fibers. As will be noted,  $C_{L0}$  equals the pure lanolin density.

### 3.2 Initial state of the bed. Lanolin breakthrough at bed exit.

An expression for  $C_L(t)$  is available if we consider the molecular diffusion of lanolin over a quiet liquid layer in contact with the solvent and a fluid-side mass transfer coefficient,  $D$  and  $k_G$ , respectively [18]. It is not difficult to show [19] that the expression of  $C_L(t)$  at bed inlet,  $z = 0$ , is:

$$C_L = C_{L0} e^{\frac{R k_G K}{D} \left( 1 - \sqrt{1 + \frac{r_0^2 k_G a K t}{R^2 (1 - \epsilon)}} \right)} \quad (9)$$

Eqn (9) indicates that  $C_{L0}$  drops rapidly to zero. In eqn (9),  $D$  is the molecular diffusivity of lanolin in the liquid,  $R$  is the final radius of the wool fiber, and  $r_0$  is the initial radius.  $k_G$  is the solvent-side mass transfer coefficient. The mass transfer group,  $Bi$ , is a modified Biot number for mass transfer,

$$\frac{Rk_GK}{D} = Bi \quad (10)$$

Now, let us see what happens at the end of the bed. In our previous work, Fullana et al. [12] showed in Fig. 13 of their article the value of the extraction yield in the fluid  $C_g$  vs time during solute breakthrough for three different velocities of extraction. The response of an extraction bed consists in a time delay, called here  $n$  later, of constant concentration, in which  $C_g$  decreases more or less rapidly depending on the solvent flowrate. After a step input pulse of solvent at bed inlet,  $z = 0$ , the shape of the breakthrough curve corresponds to an inverse step at  $z = z_T$ . The shape of the breakthrough curve can be found from eqn (9) and can be fitted to an empirical function, as we discuss next.

By performing the operations under the square root of eqn (9), it is found that,

$$C_L = C_{L0} e^{Bi(1-\sqrt{1+At})} \quad (11)$$

Where, the parameter  $A$  is,

$$A = \frac{r_0^2}{R^2} \frac{k_G a K}{1-\varepsilon} \quad (12)$$

After simplifying the exponent of eqn (11), the following expression in terms of  $Bi$ , is obtained,

$$C_L = C_{L0} e^{Bi} e^{-Bi\sqrt{1+At}} = c' e^{-\alpha\sqrt{n+t}} \quad (13)$$

Where  $c$  is a constant,  $\alpha = Bi/\sqrt{n}$ , and  $n = 1/A$ . Eqn (13) has the form of the cumulative probability of the well-known Weibull's distribution function [19]. This has a form of the type,

$$C_L = c' e^{-\alpha(n+t)^b} \quad (14)$$

In this equation,  $n + t$  is the total time, and  $b = 1/2$ . The graphical form of  $C_L$  given by eqn. (14) is the response to a unit step after time  $n + t$ . Therefore, knowing the values of  $b$  and  $\alpha$ , the Weibull expression (14) gives the breakthrough curve, that is, the decreasing sigmoid curve at bed exit after the step pulse at the inlet. Knowing the exit breakthrough curve, the extraction yield of lanolin as a function of time can be obtained during an extraction run, as we show next.

The fraction of solute extracted at time  $t$  in a run,  $X(t)$ , in terms of  $C_g(t, z_T)$  at bed exit,  $C_g(t, z = z_T)$ , is obtained with the following expression,

$$X(t) = \frac{\int_0^t C_g(t, z_T) dt}{\int_0^\infty C_g(t, z_T) dt} \quad (15)$$

In this equation, the function  $C_g$  is the concentration of solute in the extraction solvent at bed exit and time  $t$ . Let us call  $B$  the denominator of eqn (15).  $B$  represents the total solute extracted in a run, hence, it is a constant independent of time. Differentiating eqn (15) with respect to time, we find that  $C_g(t, z_T)$  is related to  $dX/dt$ , as

$$C_g(t, z_T) = B \frac{dX}{dt} \quad (16)$$

Now, if  $C_g(t, z_T)$  is approximated with the Weibull's function (eqn. (14)), the extraction yield,  $X(t)$ , is obtained by integration of  $dX/dt$  from eqn (16), as

$$X(t) = c \int_0^t e^{-\alpha(t+n)^b} dt \quad (17)$$

Where  $c$  is a constant. We then come to the following important conclusion.

When the breakthrough curve is approximated with a Weibull function (with known  $\alpha$  and  $b$ ), the fraction extracted as a function of time,  $X(t)$ , can be calculated by the integral of the Weibull function as seen in eqn (17). To simplify the adjustment of  $c$ ,  $\alpha$ ,  $b$  and  $n$ , an analytical expression of the integral in eqn. (17) is found by means of the following expression,

$$X(x) = \frac{c}{b\omega} \int_0^x x^{\frac{1}{b}-1} e^{-x} dx = \frac{c}{b\omega} \gamma\left(\frac{1}{b}, x\right) \quad (18)$$

Where  $\alpha = \omega^b$  and  $x^{1/b} = \omega t$ , and  $\gamma\left(\frac{1}{b}, x\right)$  is the incomplete Euler's gamma function with  $1/b$  argument, which might not to be integer. This function can be approximated by Taylor power series using the expression reported by Press et al. [4],

$$\gamma(m, x) = e^{-x} x^m \sum_{k=0}^{\infty} \frac{x^k}{(m+k)(m+k-1)\dots m} \quad (19)$$

The latter is very suitable for  $x \ll a$  and for arguments not necessarily integers nor positive. In this case, if time is expressed in hours, with enough terms of the series from eqn. (19) an approximation with little error can be found. Thus, applying eqn. (19) to eqn. (18) and replacing original variables, following expression can be easily found [19],

$$X(t) = ce^{-\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^b} (t-n) \sum_{k=0}^{\infty} \frac{\frac{1}{b} \left(\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^b\right)^k}{\left(\frac{1}{b}+k\right)\left(\frac{1}{b}+k-1\right)\dots\frac{1}{b}} \quad (20)$$

Where  $\alpha = \left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)\right)^b$  as the limit of the function at infinity must be equal to one. In our calculations, analytical eqn. (20) for the integral of the Weibull has been employed to obtain augmented data of the lanolin extraction curves.

### 3.3 Artificial neural networks

Neural networks are models inspired by the structure and function of biological neurons. Like those, they can recognize patterns, structure disordered data and in general learn from observation. See Gupta [13] for a modern review publication on Deep Learning in chemical engineering. A neural network is able to learn by changing the connecting weights between



neurons during the learning stage. After having learned, the knowledge is somehow stored in the weights.

In the present work, an ANN system is used to predict the extraction yield of lanolin from wool from a number of learning experimental data sets. The variables that affect the extraction rate are the essential input data. Here we used those given by Eychenne et al. [5] and were summarized in Table 1. In the present case, the input neuron layer would contain 5 units. This means that the extraction yield depends on the following 5 variables: pressure, temperature, solvent flowrate, wool compression, and time.

The effect of time is very important in the architecture of the network and how this works. A typical extraction run is dynamic in nature, because the extraction yield in a run depends on time [5]. Let us see how we could construct the net. One could add as many input neurones as there are times in an extraction run. For example, if there are 4 extraction yields at 4 different increasing times, 4 extra neurons could be used in the input layer to account for extraction yields at those times. That would be correct. However, there is a more efficient ANN to use when we have a time-varying output. The best network method for a dynamic process is to use a recurrent network, where time is taken into account. Bulsari and Palosaari [18] show why recurrent networks are best employed in their adsorption case with solute breakthrough. We implemented this type of network in our extraction case, because desorption is somehow the reverse process of adsorption, where solute breakthrough at the end of the column is desired.

### 3.4 Neural networks tools available

In our early work on neural computing [12], we employed our own programs for the design and training of the neural network. We used the backpropagation algorithm to calculate the weights during the learning stage. Modern authors, however, prefer to use optimization algorithms more efficient than backpropagation. In this work, we have used the Levenberg-Marquardt (LM) algorithm [18]. The general purpose neural network available in the Matlab toolbox is the *feedforwardnet*, that uses the LM algorithm for the weights. In our work, other Matlab methods for time-dependent outputs, based on recurrent networks, were employed, such as *NARXNET*.

### 3.5 Recurrent networks and dynamic data.

A feedforward network has no notion of order in time, and the only input it considers is the current example to which it has been exposed in the learning stage. Feedforward networks do not remember their recent past. Recurrent networks, on the other hand, take as their input not only the current input learning data that they see, but also what they have perceived previously in time. In dynamic networks, the output depends not only on the current input to the network, but also on the current or previous inputs, outputs, or previous states of the network. This allows the possibility to use neural networks for such applications where time-dependent or historical data are produced, either with randomly time-varying inputs or with ordered inputs increasing or decreasing with time. There are many examples of this; here we cite only two of them.

One example is the randomly varying stock market values over a certain period of time, in which the stock value is to be predicted. Another example is the temporal evolution of a physical variable that takes place at increasing values. This last example is our case, in which

data during an extraction run were measured over a period of time of about one hour in an experimental run. In this case, the progress of extraction increases with time. For time-varying systems, the so-called non-linear autoregressive neural network with external inputs (or *NARXNET*), is very useful when the NN Matlab toolbox is employed. A feature of *NARXNET* is that the NN can learn to predict one time series with given past values of the same time series over some previous period. The use of this NN is useful for the modeling of nonlinear dynamic systems. *NARXNET*, available in Matlab, has been used in the present work.

#### 4. Data treatment for neural networks

In paragraph 3.5, we have reasoned why we used *NARXNET* in the present work. This consists in a feedforward network that accepts two types of inputs, each multiplied by a so-called *delay*. The first type of input is an external input, these are values of previous times on which the output depends. The second type of input is a feedback input, that corresponds to a value or values of previous times of the same output. So that the output at time  $t$  depends on such previous times. Mathematically, this dependence can be expressed as,

$$Y(t) = f [Y(t - 1), Y(t - 2), \dots, Y(t - n_Y), U(t - 1), U(t - 2), \dots, U(t - n_U)] \quad (21)$$

Where  $Y$  is the output variable and the input variable of previous inputs, and  $U$  is the external variable. Then the delays applied to each input are a vector of integers (1, 2, 3....). These indicate the number of passes of previous  $Y$ s that are the inputs to the network. In other words, if the network must only consider the immediately previous one, then, the delay =1.

Instead, if the network should consider the 3 previous times, then the delay will be = (1,2,3). In this case, the network designed will be a recurrent one whose external input is the state of operation of the extractor, that is the P, T, Q and  $\rho_B$ , while the feedback input will be the values of the fraction extracted. The delay will be 1, because we want that the network learns a value of 1 previous pass.

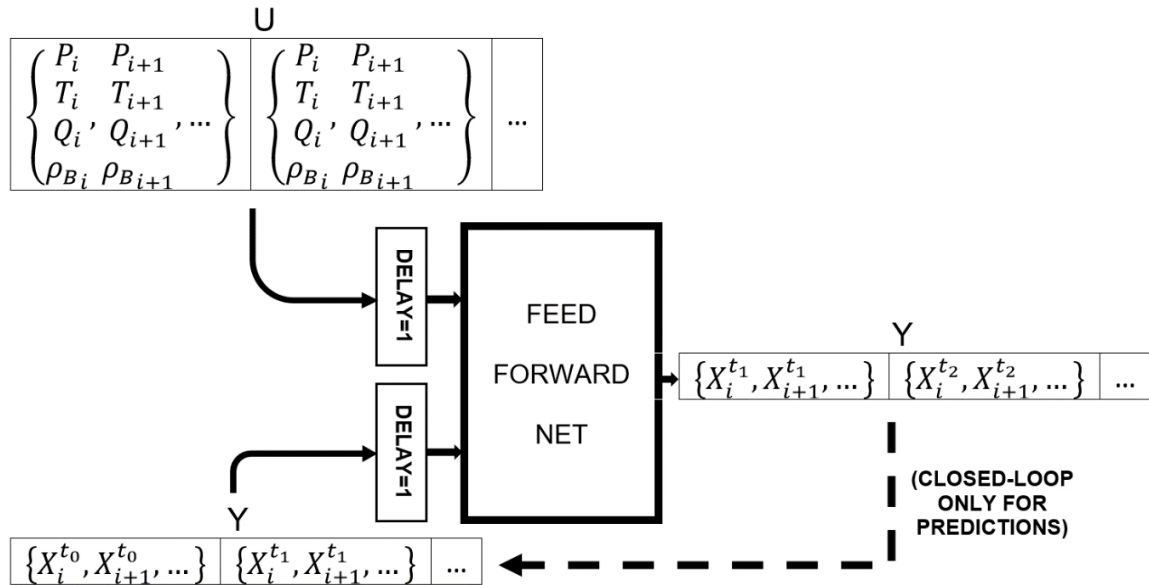
As far as training is concerned, it is very important how data are entered to the network, because in our case there are several extraction curves corresponding to different operation states. If the *NARXNET* net has to understand correctly, both the external input and the feedback is a cell array, where for each cell all operation states (for U) or extracted fractions (for Y) of a given time are contained. In such a way that the values at  $t = 0$  will correspond to the first cell, in the next cell there will be the values after 2 time steps, etc. Each cell of the external input will be formed by a matrix whose rows will consist of a different variable P, T, Q, or  $\rho_B$ , and each column will correspond to a state of operation of a different curve. Therefore, that matrix will have 4 rows, and 10 columns at most. As regards to feedback input, this will be a row vector corresponding to the extracted fraction for each different state of operation for the time of the corresponding cell, therefore it will have at most 10 columns. As regards to network architecture, we have used in principle the default values for *NARXNET*. This has one hidden layer with 10 neurons, which is the double of nodes of the input layer. This architecture will be revised later, as will be commented later, in the Results and Discussion, paragraph 4.

#### 4.1. Training data set and test data set

In order to validate the network, only 9 of the 10 states of operation are used for the training phase, that is, saving only one state (the test data set) for checking the validity of the network. The selection of the test data set is not made at random because there is a limited number of states of operation, and it is our aim that the network correctly learns the effect of each variable. For robustness, we have used for training all the combinations in which only one variable changes while the other three remain constant, so the effect of this variable can be learned properly. For the testing, we have selected a state of operation in which the evolution of its variables can be seen from other groups of data. For this reason, the state reserved for the test is as: 120 bar, 80°C, 5 kg/h and 127 kg/m<sup>3</sup>.

Once the network structure has been chosen, in order to make a prediction after the network has been trained it will be necessary to introduce the desired state of operation as an external input, together with the first extraction value that is taken as a feedback to calculate the next one. In all cases the initial value is  $X_i^{t=0} = 0$ . Then, the network works in closed-loop to predict the rest of the extracted fractions for the extraction curve at successive times, always holding constant the same external state of operation.

Fig 2 illustrates schematically how the network works.



**Fig 2.** Input, output and delays in a recurrent neural network. For predictions, not for training, the net works in closed-loop entering the previous output as the feedback input of the network.  $i, i+1, \dots$ , correspond to a different data set (or extracted fraction curve), and  $t_0, t_1, \dots$ , are the different time steps.

#### 4.2. Training parameters and prediction error

The settings for the training parameters used by Matlab are shown on Table 3.

**Table 3. Matlab training parameter values.**

Training Parameter	Value
Learning rate	0,001
Maximum iterations	Infinite
Goal (maximum mean squared error)	$10^{-20}$
Maximum time	4 min
Minimum gradient	$10^{-20}$
Maximum number of validations	Infinite

The index of performance of a NN used here is the absolute average of relative deviations, defined statistically [13] as,

$$AARD\% = \frac{1}{N} \sum_{i=1}^N \frac{|X_i^{exp} - X_i^{NN}|}{X_i^{exp}} 100 \quad (22)$$

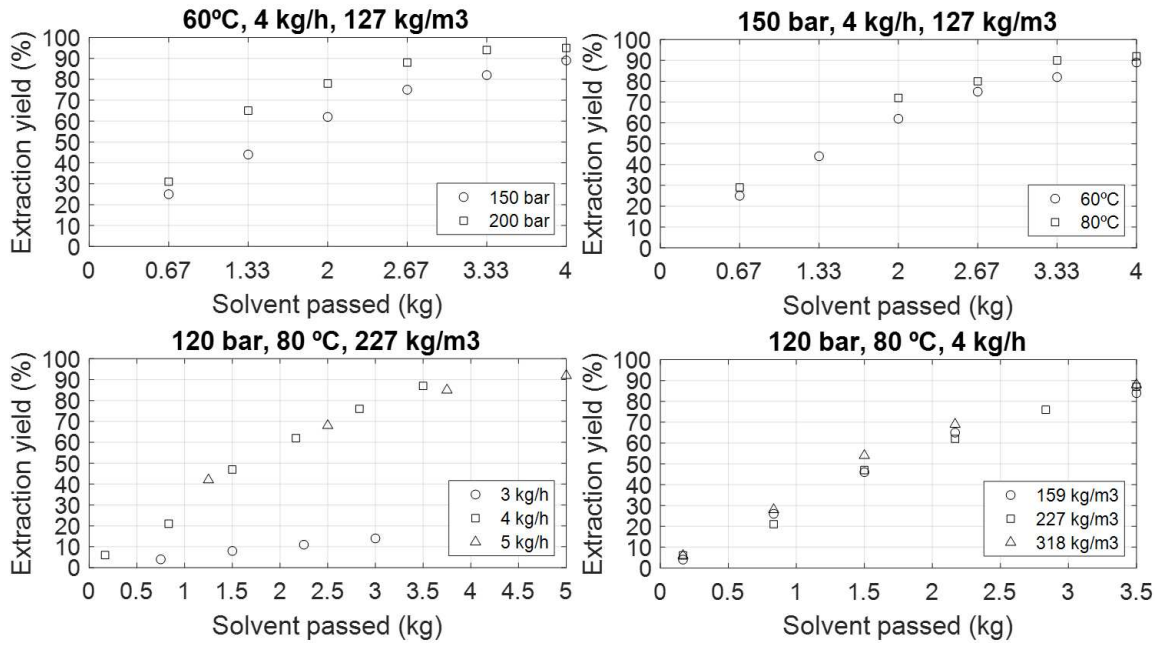
Where  $N$  is the number of data in a curve, that is, the number of time steps evaluated,

$X_i^{\text{exp}}$  is the extracted fraction regressed with the Weibull function, and  $X_i^{\text{NN}}$  is the fraction predicted by the network. As will be discussed in the Results and Discussion, for all the curves included in the training set and in the test set, their AARD% was calculated. In all cases, if this value is less than the tolerance, then the network is accepted, otherwise the training is repeated. The first step is the selection of the architecture of the network. As will be discussed later in the results, with 1 or 2 neurons in the hidden layer the result with the test is very good, but the training error increases. Conversely, a large number of neurons has the opposite effect. Clearly, there is an optimum number of neurons in the hidden layer.

## 5. Results and discussion

### 5.1. Data augmentation

The present work is based on the experimental data taken from the crude data reported in Fig 3. Only temperatures above 50°C and pressures above 100 bar have been considered. On Table 1 the different states of operation in terms of P, T, Q and  $\rho_B$ , are given. For every state, measurements of the extracted fraction were done at only 6 times (sometimes only 4), within about 1 h for each run. As will be noted, for a run the number of data is very scarce. Therefore, a method for data augmentation is necessary. As discussed in paragraph 3.1, the integral of the Weibull function can be used to regress the observed extraction values. To this purpose, an analytical integral for eqn. (19) was used. In practice, the parameters  $\alpha$  and  $b$ , for each run were obtained.



**Fig 3. Experimental data from figures 5, 6b, 8b, 9b and 10 of Eychenne et al. [5]. Data available at 120 bar, 80°C, 3 kg/h and 127 kg/m<sup>3</sup> has not been represented as it matches with the exact same values of those at same conditions and 4 kg/h.**

In order to assess that data augmentation and that the fitted extraction curves are correct, the densities of solvent and other lanolin properties are necessary; these were taken from our previous work [6].

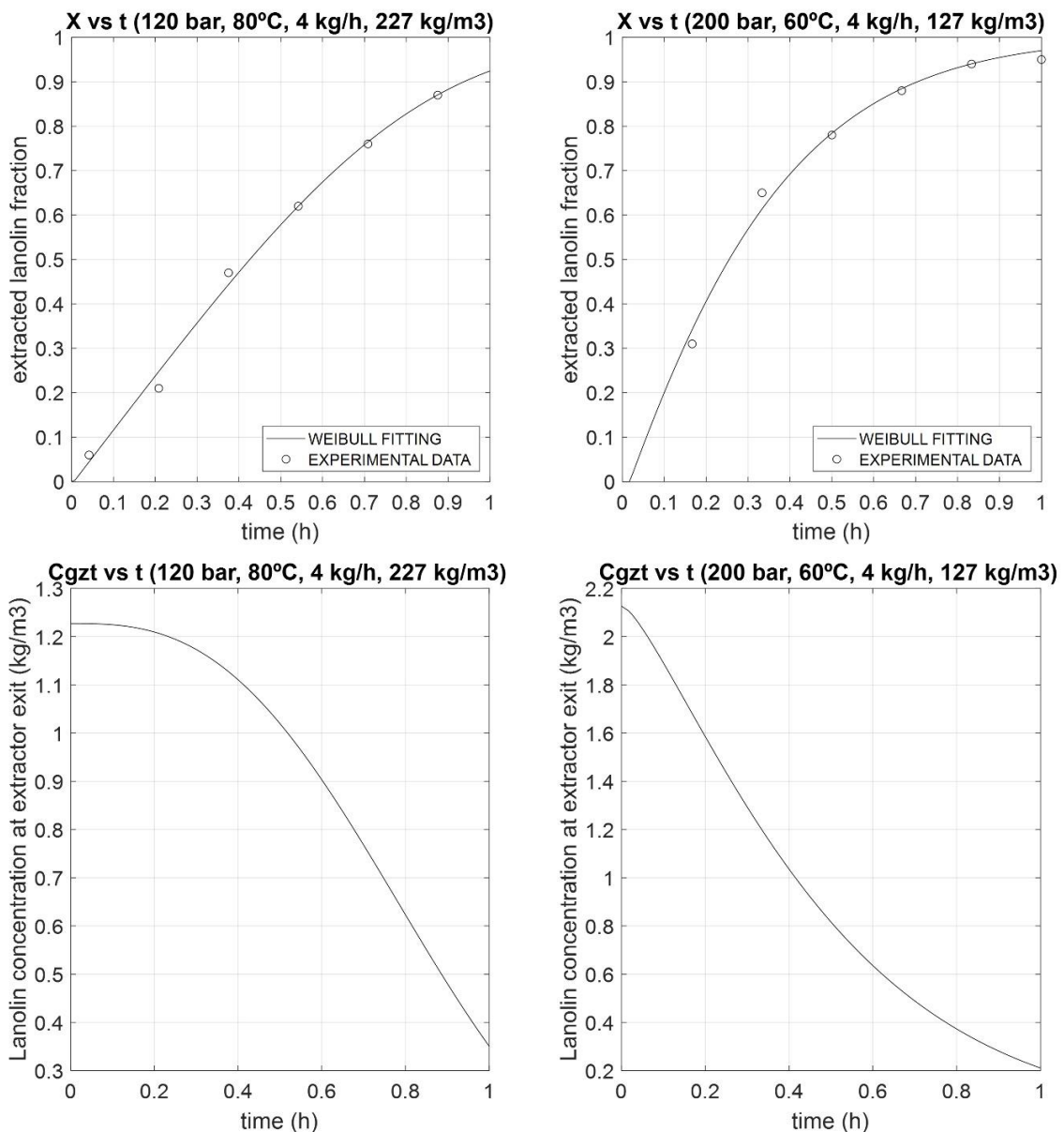
## 5.2 Fitting the experimental extraction data

The fitting procedure used in data augmentation depends on the parameters  $c$ ,  $\alpha$ ,  $b$  and  $n$ . The time  $n$  represents the time shift from which the  $X(t)$  starts to have a positive value, so that it corresponds to the point  $(n,0)$ . In order to calculate  $n$ , we apply its definition, that is,  $n$  is the time taken by the solvent to reach the end of the bed. So  $n$  can be calculated simply as,

$$n = \frac{z_T \varepsilon}{u} = \frac{z_T \pi R_B^2 \varepsilon \rho_S}{Q} \quad (23)$$



Finally, when the parameters  $c$ ,  $\alpha$ ,  $b$  and  $n$  for all the runs of Fig 3 are known, it is possible to calculate the value of  $C_g(t, z_T)$  for the experimental runs. A sample of this calculation is shown on Fig 4. It is clear that the Weibull function fits the experimental data points very well.



**Fig 4. Top curves are the fractions of lanolin extracted,  $X(t)$ , fitted with the Weibull function after data augmentation (continuous lines), compared with the measured extraction data. Bottom curves are the lanolin breakthrough at end of bed calculated by differentiation of the top curves.**

### 5.3 Designing the network. Number of nodes in the hidden layer

As noted in paragraph 4. (Data treatment for neural networks), there is a substantial difference in the prediction error in terms of the AARD% for the training data and for the test data when the number of neurons in the hidden layer is changed. Recall that the ranges of the variables P, T, Q,  $\rho_B$  are those of Table 1. Our first result is that the default value applied in NARXNET, 10 neurons, is unnecessarily large. Instead, we found that a reduced number of 2 to 5 neurons is sufficient to obtain a small AARD%. In Table 4 we show the results of the errors as a function of neurons in the hidden layer.

**Table 4. Error as AARD% for each training set as a function of nodes in hidden layer<sup>a,b</sup>**

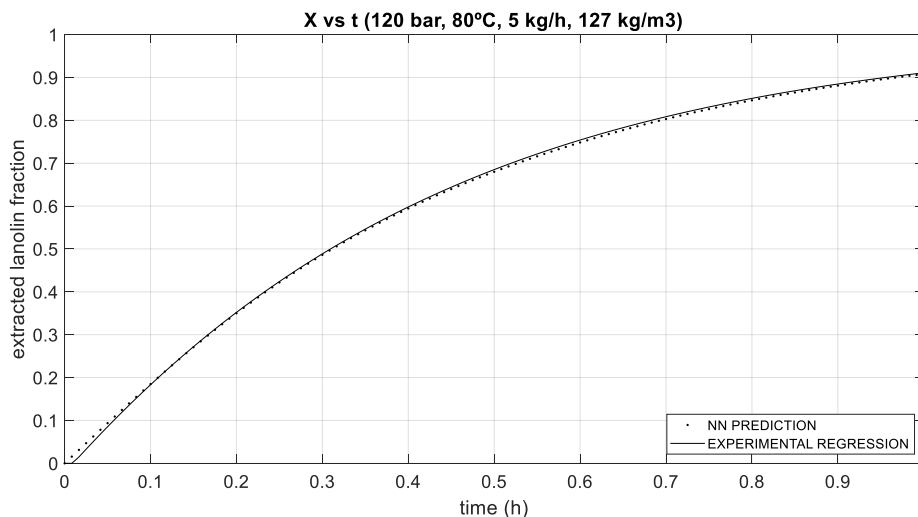
Number of neurons	Average Training Error <sup>c</sup>	Error in test data set
2	5,55 %	0,47 %
3	4,15 %	0,42 %
4	3,13 %	2,76 %
5	2,63 %	3,68 %

<sup>a</sup>With a tolerance of 10% for training sets and 5% in the test set

<sup>b</sup>Training sets include the variables P, T, Q and  $\rho_B$  of Table 1

<sup>c</sup>Average AARD% of 9 training sets

We see that the smallest error in the test set is obtained for 3 neurons. See comparison in Fig 5.



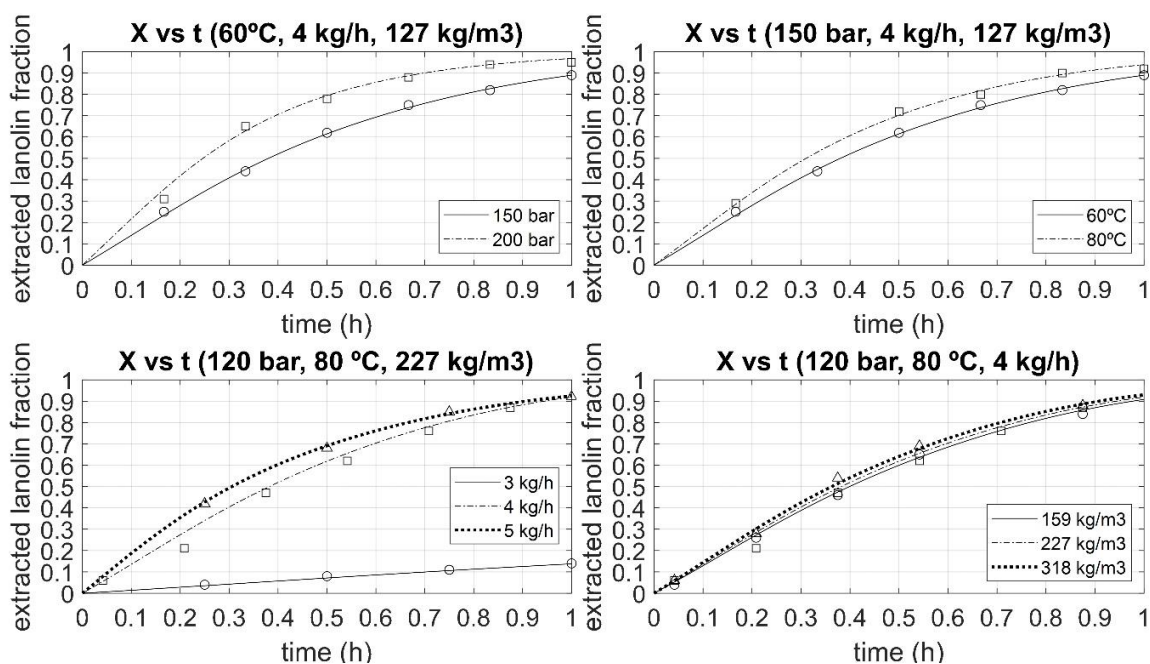
**Fig 5. Extracted fraction predicted with the test set using a trained neural network with 3 neurons in the hidden layer.**

#### 5.4 Results on the individual effects of the variables.

The merit of the neural networks developed in this work is to provide a tool that predicts the effects of the variables on the extraction process. That function is,

$$X = f(P, T, Q, \rho_B, t) \quad (24)$$

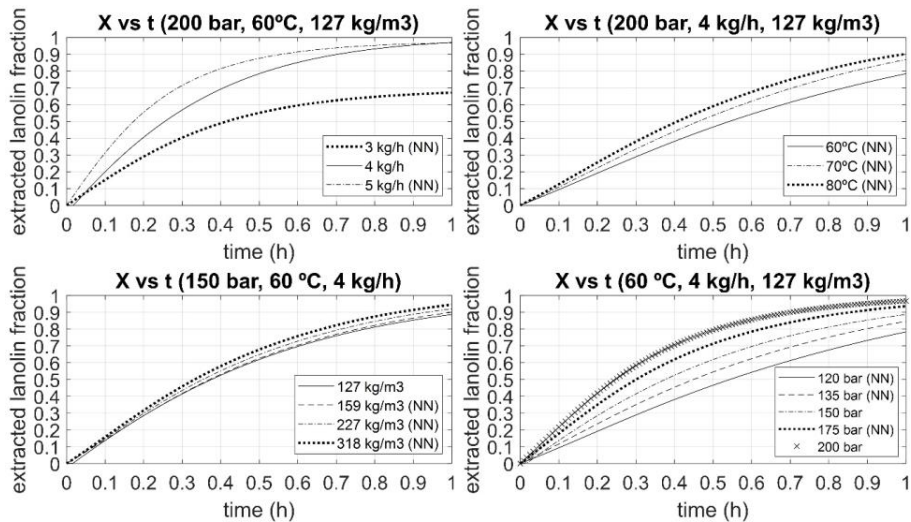
As to the effect of time, an initial condition has also been imposed, i.e., the extracted fraction at  $t = 0$  is  $X = 0$ . From then on, time is taken into account in closed-loop for as many time steps as needed to predict the curve  $X(t)$  up to the desired final time. With these conditions, the network is a function of time, and the effects of the variables can be compared with the experimental results. The results are shown on Fig 6, where the separate effects of the main variables ( $P, T, Q$  and  $\rho_B$ ) are compared with the experimental data on extracted fraction.



**Fig 6. Effect of the pressure, temperature, flow rate and packing density on the extracted fraction. Data points are the measurements [5]. Continuous lines are the neural network predictions.**

Fig 6 shows that the network prediction is generally quite satisfactory. It is observed that for a run time of 1h, the extracted fraction  $X(t)$  increases when any one of the variables increases: pressure, temperature, flowrate, or packing density. Assuming that this trend continues even

though experimental data are not available, the results shown in Fig 7 would be the expected results.



**Fig 7. Calculated effects of pressure, temperature, flow rate and packing density on the extracted fraction curves when experimental data are not available.**

It is seen on Fig 7 that in all cases the prediction of the neural network follows a trend that is consistent with the results of Fig 6. For the calculations shown on Fig 7, the network has been used not only to predict when several state variables change relative to the experiments, but also to predict values where only one variable change relative to the experiments. The conclusion is that the neural network developed seems to predict reasonably well in quite different situations.

### 5.5 Optimization problem formulation using Genetic algorithm

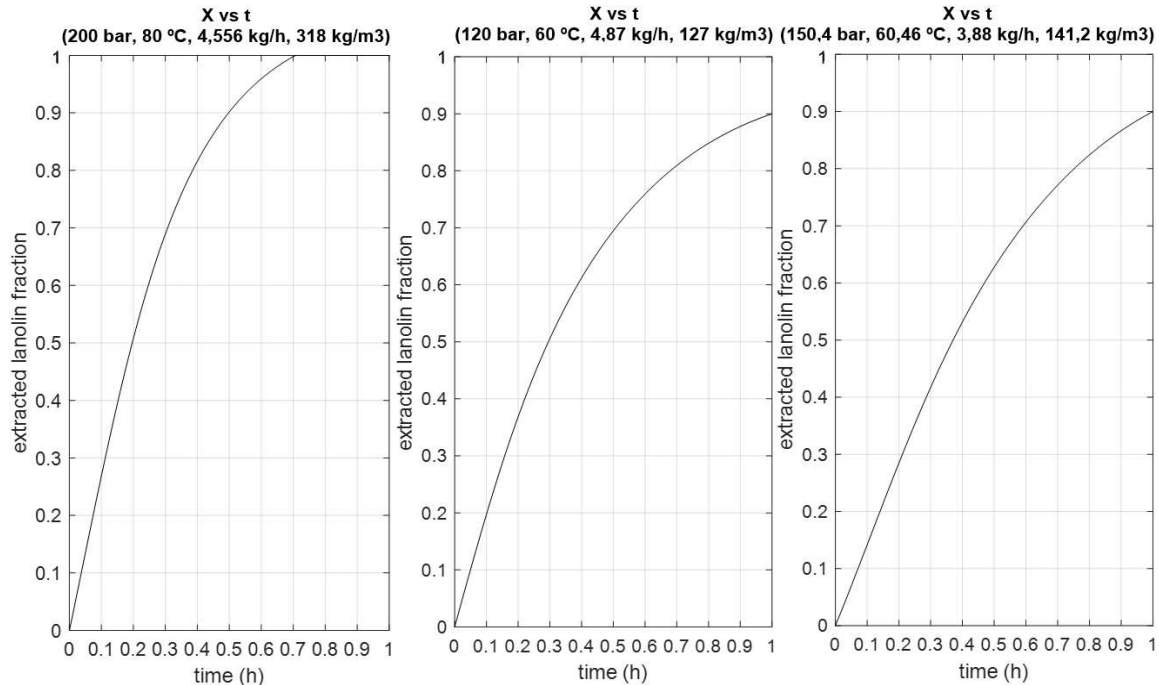
Our last finding allows to use the tool developed to find optimum operating conditions for the extraction process. There are two types of optimization problem formulations to be considered.

The first optimization problem considered can be stated as follows. For a total extraction time of 1h, find the best set of variables that maximize the extraction yield. To solve this problem we have applied the global optimization algorithm called GA (*genetic algorithm*) available in Matlab. Since once defined, the network is able to calculate the predictions quite fast, the solution to the problem proceeds very rapidly to the global optimum with a low computational cost, despite GA is known to be quite slow for cases where the objective function involves many calculations. The result of this optimization problem is readily obtained, the maximum yield is for the following set of variables: 200 bar, 80°C, 4,55 kg/h, 318 kg/m<sup>3</sup>. That maximum yield is obtained in 0,70 h (42 min).

A second optimization problem is stated as follows. Find the minimum operating conditions to obtain at least a 90% extraction yield. We also used the GA algorithm for this problem. Depending on the selection of the objective function, we could add a larger weight to a certain variables to emphasize a particular aspect of the process. For example, we can be interested in economy or in simplicity. For instance, no weight can be added, so that as a result the most important variable is solvent flowrate Q. We know, however, that the cost of increasing flowrate affects economy unfavorably. By contrast, if the variables are standardized depending on the limits of each variable, an optimum is obtained for all variables changed proportionally.

The results of both types of optimization are shown in the extraction curves of Fig 8 where the left one is the result of the first optimization, and the other two curves at right show 2 possible optima of the second optimization possibility, first without weights added and

finally with standardized data. It can be seen that those are the expected choices by inspecting the responses in Fig 7, and the effects of flowrate on  $X(t)$  shown in Fig 6.



**Fig 8.** First curve at left: extracted fraction curve for the optimum operating conditions (200 bar, 80 °C, 4,56 kg/h and 318 kg/m<sup>3</sup>) to obtain the maximum extraction yield in 1 hour. 2Other two curves at right: extracted fraction curve for the optimum operating conditions to obtain at least 90% of lanolin extracted in 1 hour. Two possible solutions are shown; left: 120 bar, 60 °C, 4,87 kg/h and 127 kg/m<sup>3</sup>; right: 150,4 bar, 60,46 °C, 3,88 kg/h and 141,2 kg/m<sup>3</sup>.

## 6. Conclusions

In this work, we have used neural network computing for modelling the behavior of the near-critical extraction of lanolin from raw wool, as a function of the variables that affect the extraction yield. These variables are: pressure, temperature, mass flowrate, packing density, and time. The recurrent neural network designed is capable to predict the extracted fraction for any set of variables, within the scope of the experiments available. Our reference experiments were those previously published by our research group.

Once more, it has been shown that it is possible to design a neural computing model that has many applications in SCF extraction, without the need to develop and solve a complex physical model of the extraction process, and without a knowledge of the laws that govern the mass transfer rate within the extraction vessel.

The potential of this network has also been discussed, showing interesting applications, as it allows, among others, to develop studies to optimize the operating variables of the extractor. A powerful global optimization algorithm, such as genetic algorithm, may be used with a low computing time cost, because once the network is trained, it works like a simple function with basic arithmetic operations. Moreover, it offers the possibility of applying weights to the variables so the desired variable is enhanced.

Another development of this work has been to employ a statistical function that allows to model the lanolin breakthrough from a packed bed of wool from which lanolin is extracted. The function used is the Weibull's cumulative probability distribution function. The integral of the Weibull function allows to regress the rather scarce extraction data and to produce a sufficient data augmentation to use in the neural algorithms. The use of the Weibull function is dictated by a physical model of extraction available to us from previous work.

**Nomenclature**

a	Mass transfer area, 1/m
A	Const. defined by eq (12), 1/s
AARD%	Absolute average of relative deviations, eqn. (22)
b	Weibull parameter,
Bi	Biot number for mass,
B	Const. eqn. (16), kg
bi	NN bias neuron
c	Const eqn. (17)
c'	Const eqn. (14)
C <sub>g</sub>	Fluid-phase concentration, kg/m <sup>3</sup>
C <sub>L</sub>	Liquid-phase concentration, kg/m <sup>3</sup>
C <sub>L0</sub>	Initial lanolin liquid, kg/m <sup>3</sup>
D	Lanolin diffusivity, m <sup>2</sup> /s
K	Henry equilibrium const.
k <sub>G</sub>	Fluid-side mass transfer coef, m/s
m	Incomplete gamma function argument, eqn. (19)
n	Time, s
P	Pressure, bar
Q	Mass flowrate, kg/s
R	Final radius, m
R <sub>B</sub>	Bed radius, m



$r_0$	Initial radius, m
$r_v$	Mass transfer rate, $\text{kg/m}^3/\text{s}$
$t$	Time, s
$T$	Temperature, $^{\circ}\text{C}$
$u$	Superficial velocity, m/s
$U$	External input in NN
$V$	Total bed volume, $\text{m}^3$
$V_L$	Liquid volume, $\text{m}^3$
$w_{ij}$	NN weights
$x$	Variable eqn. (18)
$X$	Extracted lanolin fraction,
$Y$	NN output
$z$	Length coordinate, m
$z_T$	Total length of bed, m

**Greek symbols and acronyms**

$\alpha$	Weibull parameter
$\varepsilon$	Porosity,
$\rho_B$	Packing density, kg/m <sup>3</sup>
$\rho_L$	Lanolin density, kg/m <sup>3</sup>
$\rho_S$	Solvent density, kg/m <sup>3</sup>
$\rho_W$	Wool density, kg/m <sup>3</sup>
$\omega$	Parameter eqn. (18)
ANN	Artificial neural network
LM	Levenberg-Marquardt
SCFE	Supercritical fluid extraction
NARXNET	Non-linear autoregressive with external NN

## References

- [1] Cygnarowicz-Provost, M.; King, J.W.; Marmer, W.M.; Magidman, P. Extraction of Woolgrease with Supercritical Carbon Dioxide. *J. Am. Oil Chem. Soc.*, **1994**, *71*, 222–225.
- [2] Kuo, B.S.; Kim, J.C.; Jeon, J.H.; Bae, H.K. Desorption of Wool Grease from Greasy Wool with Supercritical Carbon Dioxide, *Hwahak Konghak*, **1992**, *30*, 491-498.
- [3] Jones, F.W.; Bateup, D.R.; Dixon, S.R.; Gray, S.R. Solubility of Wool Wax in Supercritical Carbon Dioxide, *J. Supercrit. Fluids*, **1995**, *10*, 100-111.
- [4] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. Numerical Recipes in Fortran 77, Cambridge University Press, 1993.
- [5] Eychenne, V.; Sáiz, S.; Trabelsi, F.; Recasens, F. Near-critical Solvent Extraction of Wool with Modified Carbon Dioxide—Experimental results, *J. Supercrit. Fluids*, **2001**, *21*, 23-31.
- [6] Valverde, A.; Recasens, F. Extraction of Solid Lanolin from Raw Wool with Near-critical Ethanol-modified CO<sub>2</sub>—A Mass Transfer Model, *J. Supercrit. Fluids*, **2019**, *14*. 151-161
- [7] Bayona, J.M.; Erra P.; Moldovan Z.; Domínguez C.; Jover E.; Recasens F.; Larrayoz M.A., Method for Obtaining Lipid Fractions from Wool or Lanolin using Pressurized Carbon Dioxide, Patent WO 2002100990A1, Spanish Patent and Trademark Office, Madrid, filed 16.07.2005

- [8] Domínguez, C.; Jover, E.; Garde, F.; Bayona, J.M.; Erra, P. Characterization of Supercritical Fluid Extracts from Raw Wool by TLC-FID and GC-MS, *J. Am. Chem. Oil Soc.*, **2003**, *80*, 717-724.
- [9] Domínguez, C.; Erra, P.; Bayona, J.M. Physico-chemical and Dyeing Properties of Raw Wool Extracted by Pressurized CO<sub>2</sub>/Modifiers, *Textile research*, **2010**, *80*, 651-659.
- [10] López-Mesas, M.; Carrillo, F.; Gutiérrez, M.C.; Crespi, M. Alternative methods for wool wax extraction from wool scouring wastes, *Aceites y grasas*, **2007** *58* 402-407.
- [11] Puiggené, J; Larrayoz, M.A.; Recasens F. Free Liquid-to-Supercritical Fluid Mass Transfer in Packed-beds, *Chem. Eng. Sci.*, **1997**, *52*, 195-312.
- [12] Fullana, M.; Trabelsi, F.; Recasens, F. Use of Neural Net Computing for Statistical and Kinetic Modelling and Simulation of Supercritical Fluid Extractors, *Chem. Eng. Sci.*, **2000**, *55*, 79-85.
- [13] Gupta, A. Introduction to Deep Learning: Part 1, *Chem. Eng. Prog.* , **2018**, *114*, 22-29.
- [14] Bhat, N.V.; McAvoy, T.J. Use of neural nets for dynamic modelling and control of chemical process systems, *Comput. & Chem. Eng.* ,**1989**, *14*, 573-583.
- [15] Blayo, F.; Verleysen, F. Les Réseaux de Neurones Artificiels, Paris: Presses Universitaires de France, 1996.

- [16] Bulsari, A.B. (Ed.), *Neural networks for chemical engineers*, Amsterdam: Elsevier, 1995.
- [17] Fullana, M.; Trabelsi, F.; Recasens, F. Supercritical extractor design using neural-net computing, in: Bulsari A.B.; Fernández de Cañete, J.; Kalio, S. (Eds.), *Engineering Benefits from Neural networks*, Systemteknikaan seura ry, Finland, Turku, 1998.
- [18] Bulsari, A.B.; Palosaari, S. Application of Neural Networks for System Identification of an Adsorption Column, *Neural Comput. & Applic.*, **1993** 1 160-165.
- [19] A. Valverde, PhD thesis, to be published.
- [20] NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>, date.



### **3.4. Mathematical Modelling of Supercritical Fluid Extraction of Liquid Lanolin from Raw Wool. Solubility and Mass Transfer Rate**

#### **Parameters**

(Submitted to Chemical Engineering Research and Design)

As a closure for the extraction of liquid lanolin in supercritical ethanol-modified CO<sub>2</sub> problem, a new physical model for the same operation conditions discussed in the previous paper is presented. In the above paper, a mathematical model was avoided by the complexity of diffusion phenomena inside the lanolin layer, leading to a neural network model design. In here, this problem is overcome by an assumption that simplified the model, avoiding this way the need of a theoretical radial profile of the lanolin concentration in the liquid phase. This hypothesis was considered by the good results reported by Puiggené et al. [83], and consisted on assuming an average value  $C_L$  of the lanolin concentration profile in the liquid phase, founded on the idea that a thin lanolin layer would probably have a much planar concentration profile. This assumption led to two coupled partial differential equations, namely the resulting from the dispersed flow balance and the resulting from the matching of lanolin gradients from liquid and solvent phase. It is possible to solve the latter by applying the Laplace transform first, obtaining an explicit expression for  $C_L$  which depends on the integral of the lanolin concentration in the solvent phase  $C_g$ . Setting the initial condition  $C_g(0,0) = 0$ , the obtained expression could be numerically integrated in order to find  $C_L$  time profile, the starting point for solving the first PDE. Hereby, finite differences method was used to solve the system, calculating the numerical integration of the previous  $C_g$  profile at each iteration.

The mathematical methods explored in this chapter, namely Laplace transform and finite differences, whether not being so interesting for their lack of novelty, resulted in novel mass transfer parameters, Henry constants and solubility values for lanolin supercritical extraction at different temperatures, pressures, mass flowrates and packing densities. Moreover, those values found by global optimization using Genetic Algorithm, agreed very well with those obtained in previous works and those from literature. For instance, mass transfer parameters were coherent with those reported by Puiggené et al. [83], King et al. [149] for rapeseed oil extraction on similar conditions, and Valverde et al. [104] for solid lanolin. Operation conditions effect on mass transfer parameter is also discussed, decreasing with increasing pressure (meeting Brunner [16] results), agreeing with Eychenne et al. [66] in the existence of an optimal packing density, increasing markedly with increasing flow rate and showing almost non effect of temperature. Solubilities found also agree with those reported by Eychenne et al. [66] and Henry constants do the same with those obtained by Weibull distribution modelling in the previous Valverde et al. [62] work. In this sense, this new work validate the neural network predictions, showing almost the same extraction curves.

The paper resulting from this work is under review since 8<sup>th</sup> of January (2020) in *Chemical Engineering Research and Design*, a 1<sup>st</sup> quartile journal in Chemical Engineering area, IF (2018) 3,073 [82].



# Mathematical modelling of supercritical fluid extraction of liquid lanoline from raw wool. Solubility and mass transfer rate parameters

Abel Valverde\*, Jesús Alvarez-Florez, Francesc Recasens  
Chemical Engineering Dept., Diagonal 647, Universitat Politècnica de  
Catalunya, Barcelona, Spain

\*Corresponding author

**Keywords:** Extraction, lanoline, raw wool, solubility, near-critical CO<sub>2</sub>, mass-transfer.

## Abstract

A new mathematical model is presented for the supercritical fluid extraction of lanoline from wool using near-critical ethanol-modified CO<sub>2</sub>, using our previous experimental data. The model is intended to account for the extraction of lanoline at conditions well above its melting point (60-80°C) and pressures up to 150 bar. The model parameters are a Henry-type fluid-to-liquid partition coefficient for lanoline,  $K = C_g/C_L$ , and a fluid-side mass-transfer coefficient,  $k_G$ . For  $Re \sim 1$ ,  $K$  is independent of velocity and wool packing density, but increases with pressure ( $K = 4-15$ ).  $k_G$  is found to be independent of temperature; it increases with velocity, decreases with pressure, and increases with wool packing density. The values found are  $k_G = 5.66 \times 10^{-6}$  m/s (at 70 bar) and  $k_G = 1.51 \times 10^{-6}$  m/s (at 150 bar). The predictions of our physical model agree with the results of a recently published empirical neural computing model.

**Nomenclature.** See end of manuscript.

## 1. Introduction

Supercritical fluids (SCF) have been known since the late nineteenth century. But the modern impulse of SCF reappeared in the 1970s in connection with the interest in natural products and the processes for treating valuable substances, such as coffee and tea decaffeination, hops, extraction of flavours and fragrances and pharmaceutical separation methods. SCF is now considered a mature technology. Information about SCF technology can be found in recognised texts dealing with the physical-chemical properties and applications of SCF (McHugh and Krukoniš, 1986; King and Bott, 1993; Brunner, 1994; Clifford, 1999; Arai et al., 2002).

Lanoline is a highly valued wool grease secreted by sheep and is widely used in the cosmetic and pharmaceutical industries. Early attempts to extract lanoline from raw wool with supercritical carbon dioxide date back to the 1990s, and are the work of Koo et al., 1992, Cygnarowicz-Provost et al., 1994, and New Zealand researchers (Jones et al., 1995). In all cases, pure compressed CO<sub>2</sub> was employed. Bayona et al. (2005) employed modified CO<sub>2</sub> to extract lanoline from raw wool, developing a process for the fractionation of lanoline from raw wool or technical lanoline. One of the first experimental systematic studies was that published by our group (Eychenne et al., 2001), which is the experimental basis of the present work. In this work, CO<sub>2</sub> modified with 5% ethanol was used under near-critical conditions, allowing operation at lower pressure. Extraction yields higher than 90% were obtained in about 1-2 h, which were not possible with pure CO<sub>2</sub>. Recently, we published an article dealing with modelling the extraction of solid lanoline from wool using modified CO<sub>2</sub> (Valverde and Recasens, 2019) at a single temperature, 30°C, and pressures up to 150 bar. In these conditions, the lanoline is a solid (melting point. 38-44 °C) and the fluid is a liquid. Extraction rates

were interpreted using the gas-solid shrinking-core concept for non-catalytic reactions (Fogler, 2006) applied to extraction. In this case, the solid lanoline covering the wool fibres is in direct contact with the fluid and is progressively dissolved by the flowing solvent. Lanoline is known to consist of two fractions. In the Valverde and Recasens (2019) model the two lanoline fractions are assumed to be deposited one over the other on the wool fibre, as is suggested by the empirical results obtained in the fractionation process developed in a patent (Bayona et al., 2005) and the chemical characterization of the fractions (Domínguez et al., 2003a, 2003b; López-Mesas et al., 2007).

The present work is based also on the results of the experiments of Eychenne et al., (2001), but now the modelling work is extended to wider ranges of temperature, pressure and other variables that affect the extraction rate. As can be seen in Table 1, lanoline is extracted from a liquid to a near-critical fluid at temperatures of 60 to 80°C, well above the lanoline melting point (38-42°C), in all the extraction runs. At these temperatures, the conservation equations for lanoline become complicated by the fact that there is a multicomponent liquid phase, which involves an additional mass transfer resistance compared with the extraction runs at temperatures where lanoline is a pure solid. For temperatures above 30°C, the lanoline diffusion equations from the liquid up to the liquid-fluid interface mean that solving the model is quite complex. To overcome these difficulties, Valverde et al. (2019) employed neural computing to solve the problem of characterising the effects of the variables. The neural network approach allowed us to accurately predict packed bed extraction yield as a function of time, for given values of pressure, temperature, solvent flow rate, and wool packing density. The average error for the neural network predictions was less than  $\pm 0,42\%$ . Thus, the prediction using neural networks is very accurate.

However, one shortcoming of neural computing is that it cannot provide the values of physically meaningful model parameters, such as the mass transfer coefficient and solubility of the solute as a function of the process variables. This is a problem for the design and simulation of packed beds from first principles, which requires data on equilibrium solubility and mass transfer rate parameters. (King and Catchpole, 1993).

The purpose of the present work is to solve the mass conservation equations for lanoline extraction from raw wool in conditions where lanoline is a liquid. Our aim is to determine the physical parameter values that characterise the fundamental physical model. A further purpose of this work is to compare the results of the present method with the prediction of lanoline extracted fractions using neural networks, already available. In both cases, our previous experimental extraction runs (Eychenne et al, 2001) were the basis of our models. Our work offers a complete mathematical model for the design and operation of near-critical solvent extraction of lanoline from wool.

## **2. Experimental background**

The experimental study behind this work was published elsewhere (Eychenne et al., 2001); the reader should refer to it for a complete description. Here, we report only a few significant features. A Separex 200 unit was used for the high-pressure extraction. The laboratory extractor unit is shown in Fig 1, where the elements of the setup are described. A summary of the scope of the measurements, features of the extractor vessel and the wool stock used are given in Table 1.

**Table 1. Settings for the extraction runs of raw wool. Extractor and wool fibre properties**

Temperature	60, 80°C
Pressure	70, 120, 150 bar
Solvent* mass flow rate	3, 4 and 5 kg/h
Solvent passed	up to 5 kg
Total extraction time	1h approx.
Wool packing densities	127, 159, 227 and 318 kg/m <sup>3</sup>
<b>Extractor vessel</b>	
Shape	Cylindrical
Material	Stainless steel AISI 316L
Interior dimensions	145 mm x 30 mm (H x D)
Volume	100 cm <sup>3</sup>
Internal cross section	706.8 mm <sup>2</sup>
<b>Wool stock fibres</b>	
Wool loading in extractor	13 g
Wool composition	60-65% wool proper, 10-20% wax and proteins, 10% soluble stains (salts), 1-20% soil and vegetable matter
Fibre geometry	Cylindrical
Fibre length (approx.)	15 cm
Initial fibre radius, $r_0$	$\approx 8.3 \mu\text{m}$
Final fibre radius, R	20 $\mu\text{m}$
Lanoline content	20 % wt
Types of lanoline	Two fractions (external and internal)
Lanoline melting point	38 – 44 °C
Equivalent (Sauter) particle size of wool fibres	$3.2 \times 10^{-2} \text{ mm}$

\* Solvent is 95% wt. CO<sub>2</sub> – 5% ethanol.

The runs of the extracted lanoline fraction as a function of time have been taken from the Figs 5, 6b,8b, 9b and 10, belonging to the article by Eychenne et al (2001). Only the supercritical operating data have been used in the present work, that is, for temperatures above 50°C and pressures above 100 bar. The PT envelope for the solvent mixture we know that the the critical point of the solvent 5% Ethanol-95% CO<sub>2</sub> is about 50°C and 95 bar; at 30°C and pressures above 70 bar the fluid is a liquid. For T below 38-44°C, lanoline is a solid. See Valverde and Recasens (2019).

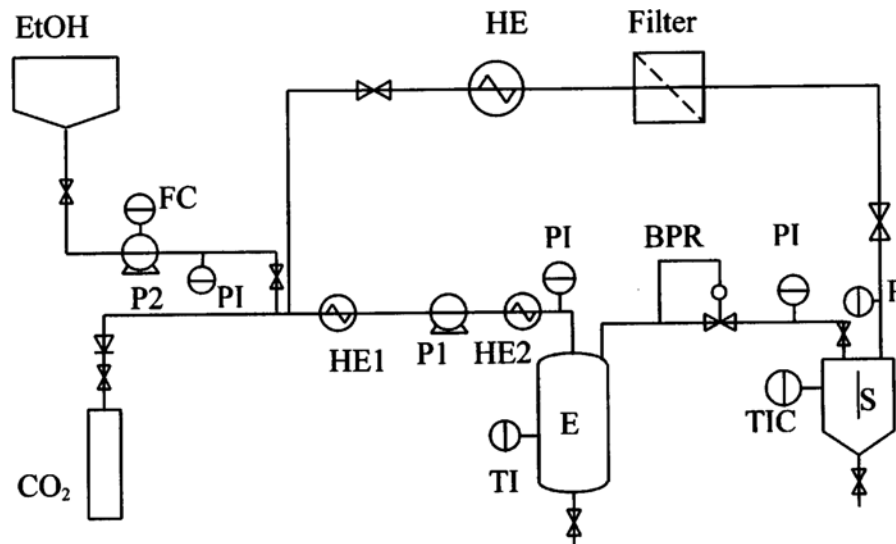


Fig 1. Separex 200unit flow diagram. P1, Milton-Royal CO<sub>2</sub> pump; P2, Pula-Feeder ethanol pump; HE1, HE2, HE3, heat exchangers; BPR back-pressure control valve; E, Separex 200 extractor vessel; S, Separex cyclone separator; PI, pressure indicators; TIC temperature indicating controller; FC, flow controller

The physical properties of the materials are required in order to solve the mathematical model for extraction, i.e., the solvent densities and the densities of lanoline and the wool fibre. These are taken from Valverde and Recasens (2019). It is worth mentioning that the total extractor length,  $z_T$ , and the void fraction of the wool,  $\epsilon$ , both depend on the operating packing density of the wool bed,  $\rho_B$ , because the total load of wool in the extractor vessel is the same in all runs (see Table 1). In order to change the packing density, the length  $z_T$  is decreased by lowering the threaded upper head of the extractor vessel. So, when the length  $z_T$  is decreased, the wool packing density increases. The relationship between packing density, bed porosity and extraction vessel length can be calculated. See Table 2.

Table 2. Wool packing density - bed porosity - extractor length relationship

Bulk density $\rho_B$ (kg/m <sup>3</sup> )	Bed porosity $\epsilon$ (m <sup>3</sup> void/m <sup>3</sup> bed)	Extractor height $z_T$ (m)
127	0.90	0.145
159	0.88	0.116
227	0.83	0.081
318	0.76	0.058

### 3. Theory

#### 3.1 Conservation equations.

Consider a bed of packed wool from which the lanoline deposited on the fibers is extracted by the flowing solvent. At  $t = 0$ , a step input of lanoline-free solvent is introduced in the bed inlet at  $z = 0$ . At the same time, the lanoline-saturated fluid is removed at the other end of the bed at  $z = z_T$ . Neglecting pressure drop, the solvent passes through the bed in plug flow at constant pressure and temperature. The rate of extraction depends on two factors: the solubility of lanoline in the solvent and a fluid-side mass transfer coefficient. Two assumptions are made:

- 1) Since the amount of lanoline on the wool is small, the bed void fraction,  $\epsilon$ , is about constant and equal to that of packed wool; and
- 2) Since the layer of liquid lanoline covering the fibers is very thin, the total mass transfer area,  $a$ , is considered constant; therefore, the holdup of liquid in the bed,  $\epsilon_L$ , is taken as equal to its average value.

With these assumptions, the balance equations for lanoline in the flowing solvent and in the liquid phase, respectively, are

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = r_v \quad (1)$$

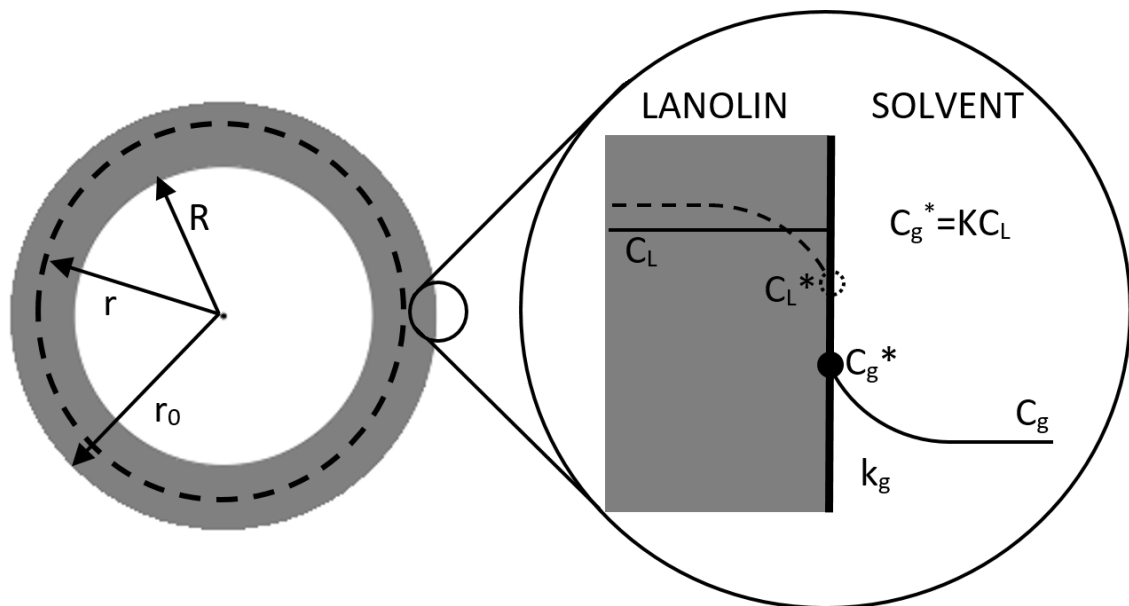
$$-\frac{\partial (V_L C_L)}{\partial t} = V k_G a (C_g^* - C_g) \quad (2)$$

Where  $r_v$  is the volumetric mass transfer rate at position  $z$  and time  $t$ , and  $k_G$  is the fluid-side mass transfer coefficient. In eqn. (2),  $C_g^*$  is the lanoline concentration in the solvent side.  $C_g^*$  is related to the concentration of lanoline in the bulk liquid,  $C_L$ ,

through a Henry-type equilibrium constant,  $K$ , as is usually the case in high-pressure studies. This is

$$C_g^* = KC_L \quad (3)$$

Fig 2 shows the initial and final liquid side lanoline radius, and the equilibrium between  $C_L$  and  $C_g^*$ , that is, between the lanoline concentration in the bulk liquid and that in the gas phase.



**Fig 2. Initial and final liquid lanoline on a wool fibre, and liquid-solvent mass transfer gradients. Note that  $C_L$  is the lanoline in the bulk liquid (see text).**

In order to develop the theoretical model, it is further assumed that the concentration of lanoline in the liquid phase has an average value  $C_L$  of the lanoline concentration profile in the liquid layer (see the right of Fig 2). That hypothesis was already made by Puiggené et al. (1997) with very good results. These authors were able to obtain an analytical expression for the fluid concentration at bed exit,  $C_g(t, z_T)$ , for the case of the liquid-to-SCF mass-transfer in a packed bed on inert particles.



Note that  $C_L$  decreases with increasing  $z$ , because lanoline dissolves in the solvent for increasing values of  $z$ . Using the following relationship

$$V_L = V \epsilon_L \quad (4)$$

together with the gas and liquid holdups,  $\epsilon$  and  $\epsilon_L$ , and the Henry constant,  $K$ , eqns. (1) and (2), become

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = K k_G a (C_L - \frac{C_g}{K}) \quad (5)$$

$$-\frac{\partial C_L}{\partial t} = \frac{K k_G a}{\epsilon_L} (C_L - \frac{C_g}{K}) \quad (6)$$

Where  $\epsilon_L$  is the liquid holdup, taken as an average value during the extraction process. If the bed is initially loaded with new wool, and the inlet solvent is lanoline-free, the boundary and initial conditions for the system of eqns. (5) and (6), are

$$C_g(t, z = 0) = C_g(t = 0, z) = 0 \quad (7)$$

$$C_L(t = 0, z) = C_{L0} \quad (8)$$

Where  $C_{L0}$  is the initial concentration of lanoline covering the fibers. As will be noted,  $C_{L0}$  equals the pure lanoline density.

### 3.2 Time evolution of lanoline concentration in the gas phase at bed inlet.

A description of the bed conditions follows from the start of the extraction. First, we study what happens at bed inlet, and later we look at bed outlet. When the solvent enters the bed at  $t = 0$  and  $z = 0$ ,  $C_g$  is  $C_g = 0$  in eqn. (6), so that the initial rate of change of  $C_L$  with time at bed inlet, will be

$$-\frac{\partial C_L}{\partial t} = \frac{K k_G a}{\epsilon_L} C_L \quad (9)$$

Integration of eqn. (9) from  $t = 0$ , gives the following expression for  $C_L$ , which accounts for the value of  $C_L$  near  $z = 0$ ,

$$C_L = C_{L0} e^{-\frac{K k_G a}{\epsilon_L} t} \quad (10)$$

This expression indicates that at the beginning of extraction,  $C_L(t)$  near  $z = 0$ , rapidly drops exponentially from  $C_{L0}$  to zero. Now, if  $C_L$  given by eqn. (10) is substituted in eqn. (5), the following expression is obtained:

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = k_G a (K C_{L0} e^{-\frac{K k_G a}{\epsilon_L} t} - C_g) \quad (11)$$

Using the initial conditions, eqns. (7) (8), eqn. (11) can be written as

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = K k_G a C_{L0} e^{-\frac{K k_G a}{\epsilon_L} t} \quad (12)$$

Using this expression and the initial conditions, eqns. (7) and (8), we can begin to solve the partial differential equations with the method of finite differences, advancing one step forward,  $\Delta z$ , and progressing to the end of the bed. See paragraph 4.1 below.

### 3.3 Time evolution of lanoline concentration in the gas phase at any point of the bed.

For the purpose of solving the conservation equations (5) and (6) at a point within the bed, it is necessary to solve eqn. (6) in order to find the value of  $C_L$  at any  $z$ , or  $C_L(z)$ , and then apply a finite difference scheme up to the end of the bed. Eqn. (6) is readily solved by the Laplace transform method, as summarized next:

$$-\mathcal{L}\left\{\frac{\partial C_L}{\partial t}\right\} = \frac{k_G a}{\epsilon_L} \mathcal{L}\{K C_L(s) - C_g(s)\} \quad (13)$$

$$s C_L(s) - C_{L0} = -\frac{k_G a}{\epsilon_L} (K C_L(s) - C_g(s)) \quad (14)$$

$$s C_L(s) + \frac{K k_G a}{\epsilon_L} C_L(s) = \frac{k_G a}{\epsilon_L} C_g(s) + C_{L0} \quad (15)$$

$$\left(s + \frac{K k_G a}{\epsilon_L}\right) C_L(s) = \frac{k_G a}{\epsilon_L} C_g(s) + C_{L0} \quad (16)$$

$$C_L(s) = \frac{\frac{k_G a}{\epsilon_L}}{\left(s + \frac{K k_G a}{\epsilon_L}\right)} C_g(s) + \frac{C_{L0}}{\left(s + \frac{K k_G a}{\epsilon_L}\right)} \quad (17)$$

$$\mathcal{L}^{-1}\{C_L(s)\} = \mathcal{L}^{-1}\left\{\frac{\frac{k_G a}{\epsilon_L}}{\left(s + \frac{Kk_G a}{\epsilon_L}\right)} C_g(s)\right\} + \mathcal{L}^{-1}\left\{\frac{C_{L0}}{\left(s + \frac{Kk_G a}{\epsilon_L}\right)}\right\} \quad (18)$$

Applying the convolution theorem to eqn. (18), it is inverted to give

$$C_L = \frac{Kk_G a}{\epsilon_L} \int_0^t C_g(\tau) e^{-\frac{Kk_G a}{\epsilon_L}(t-\tau)} d\tau + C_{L0} e^{-\frac{Kk_G a}{\epsilon_L} t} \quad (19)$$

$$C_L = \left( \frac{k_G a}{\epsilon_L} \int_0^t C_g(\tau) e^{\frac{Kk_G a}{\epsilon_L} \tau} d\tau + C_{L0} \right) e^{-\frac{Kk_G a}{\epsilon_L} t} \quad (20)$$

In brief, after finding the first step from the bed inlet using finite differences, it is possible to obtain the integral of eqn. (20) numerically (for example, by trapezoids under the  $C_g(t)$  curve). With this in hand, the profile of  $C_L(t)$  is readily obtained from eqn. (20). Once this is calculated, eqn. (5) is used again to advance a step in  $z$ . The process is then repeated, up to the end of the bed. As pointed out before, Puiggené et al. (1997) obtained an analytical solution using the incomplete gamma functions in the form of a series. Puiggené's method is significantly more complex compared with the rather simple numerical integral just obtained, with a comparable error.

Fig 3 shows the three positions of the extraction front along the bed as a function of time, as we have just discussed. The beginning of the breakthrough curve at bed exit is also represented on the right.

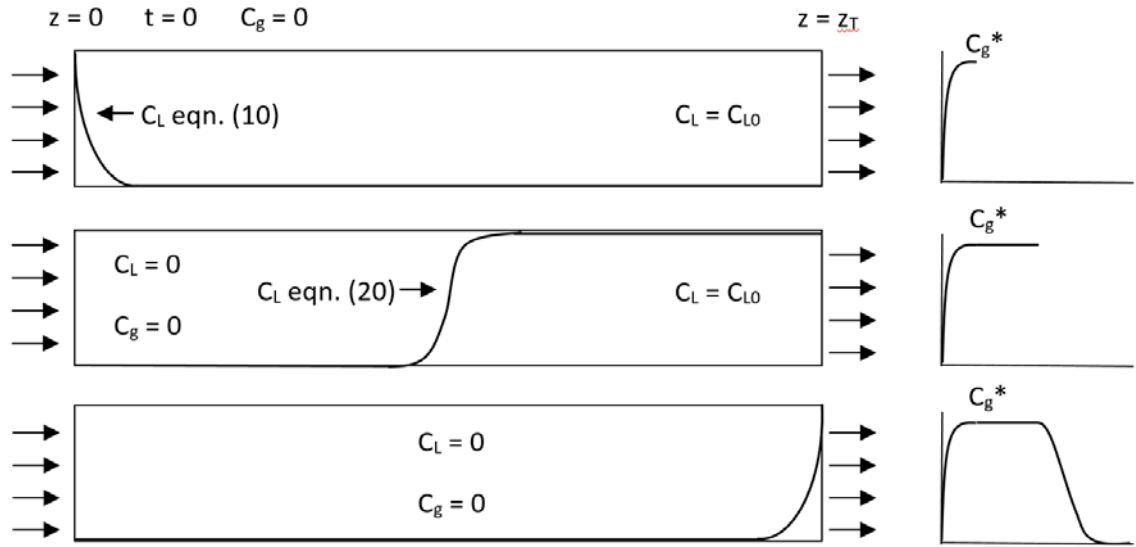


Fig 3. Extraction front through the extractor. A step input of pure solvent enters at  $z = 0$  on the left. Top: at the initial moment; middle: at half extraction, bottom: near the end of the extraction.

## 4. Method of solution

### 4.1 Integration of the model eqns. (1) and (2)

The method of finite differences is used to integrate the system of equations (1) and (2).

To do that let us write eqn. (1) in the form of finite increments, as

$$\epsilon \frac{\Delta C_g}{\Delta t} + u \frac{\Delta C_g}{\Delta z} = r_v \quad (21)$$

For the time derivative, the central difference for  $\Delta C_g$  is used. For the space derivative, the forward difference is used, that is, the difference between the value of  $C_g$  at the last  $z$  and the value at the next  $z$ , which is unknown. So, eqn. (21) is written as follows:

$$\epsilon \frac{C_g(z_i, t_i + h_t) - C_g(z_i, t_i - h_t)}{2h_t} + u \frac{C_g(z_i + h_z, t_i) - C_g(z_i, t_i)}{h_z} = r_v(z_i, t_i) \quad (22)$$

Where  $h_t$  is the time step size,  $\Delta t$ , and  $h_z$  is the space step size,  $\Delta z$ . To calculate  $r_v(z_i, t_i)$ , the following expression, based on eqns. (5) and (20), is used:

$$r_v(z_i, t_i) = k_G a \left( \left( \frac{K k_G a}{\epsilon_L} \int_0^{t_i} C_g(z_i, \tau) e^{\frac{K k_G a}{\epsilon_L} \tau} d\tau + K C_{L0} \right) e^{-\frac{K k_G a}{\epsilon_L} t_i} - C_g(z_i, t_i) \right) \quad (23)$$

Now, for a given value of  $z = z_i$ , the terms that appear in eqn. (22) are all known except for  $C_g(z_i + h_z, t_i)$ , which is unknown. Once eqn. (23) is solved with a known value of  $C_g(z_i, t_i)$ , the obtained  $r_v(z_i, t_i)$  value is then used to find a new  $C_g(z_i + h_z, t_i)$  at the next  $z_i$ , isolating it from eqn. (22). Then, the procedure is repeated up to the end of the bed.

#### 4.2 Parameter optimisation

Two model parameters are subject to optimisation. These are: the Henry constant for lanoline dissolution in the solvent,  $K$ , and the volumetric mass transfer coefficient,  $k_{GA}$ . The global optimization algorithm *Genetic algorithm (GA)*, available in Matlab, was used to obtain these values. The objective function to minimise was the *Mean Squared Error (MSE)*, which involves the squared differences between the observed extracted fractions of lanoline, and the fractions predicted by the model for the same times. We chose this type of function, that is, without denominator, because of the absence of experimental or theoretical ones, which tend to overvalue the sum of small denominators and to undervalue those values close to unity. The expression of MSE is

$$MSE = \frac{1}{N} \sum_{i=1}^N (X_i^{exp} - X_i^{theor})^2 \quad (24)$$

The values of the extracted fractions,  $X(t)$ , used in eqn. (24) were calculated from the extraction runs using the following expression:

$$X(t) = \frac{\int_0^t C_g(t, z_T) dt}{\int_0^{\infty} C_g(t, z_T) dt} \quad (25)$$

## 5. Results and discussion

### 5.1 Results on parameters $k_{GA}$ and $K$ . Estimation errors.

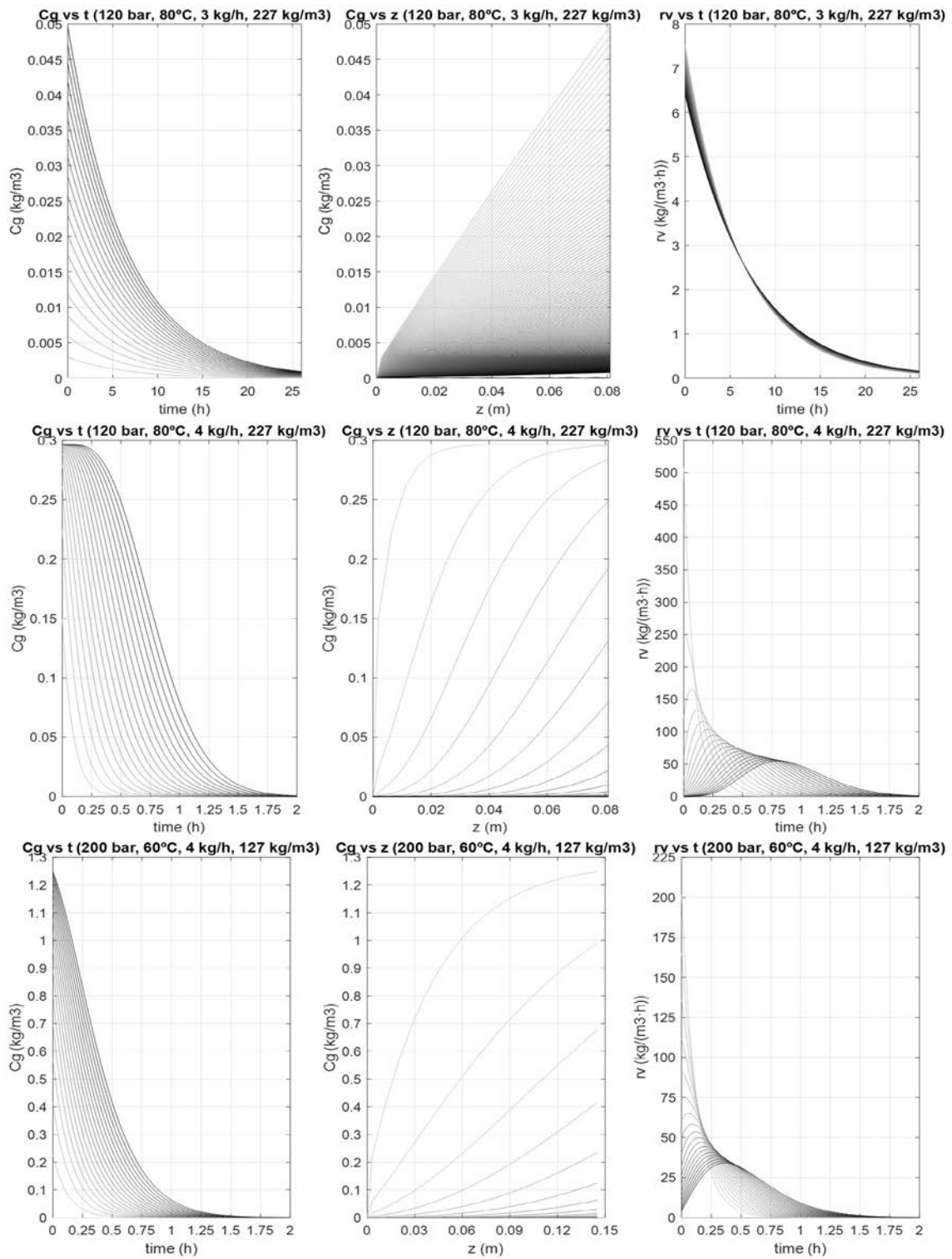
The integration of eqns. (1) and (2) using the method of finite differences, gives the lanoline concentration in the gas phase and the extraction rates,  $r_v$ , at different time  $t$  and position in the bed  $z$ . These values are represented in Fig 4. After performing the integrals of the curves  $C_g$  vs  $t$  of eqn. (25) up to the end of the bed, the fractions of lanoline extracted as a function of time are calculated. The resulting optimal parameters obtained by the GA optimisation are given in Table 3. This table shows the values of the AARD% (Average absolute relative deviation). This index is a statistical standard error indicator, better than the MSE (Fullana et al., 2000). The expression of the AARD% is

$$AARD\% = \frac{1}{N} \sum_{i=1}^N \frac{|X_i^{exp} - X_i^{NN}|}{X_i^{exp}} 100 \quad (26)$$

This expression explains why this index has not been used directly in the GA optimisation.

**Table 3. Values of model parameters obtained in this work,  $k_{GA}$  and  $K$ .**

<b>P (bar)</b>	<b>120</b>						<b>150</b>		<b>200</b>	
<b>T (°C)</b>	<b>80</b>						<b>60</b>	<b>80</b>	<b>60</b>	
<b>Q (kg/h)</b>	<b>3</b>		<b>4</b>			<b>5</b>		<b>4</b>		<b>4</b>
<b><math>\rho_B</math> (kg/m<sup>3</sup>)</b>	127	227	159	227	318	127	227	127	127	127
<b><math>k_{ga}</math> (h<sup>-1</sup>)</b>	22.6	22.6	743.7	1706.5	1348.9	899.3	970.5	171.9	191.8	157.9
<b><math>K \times 10^4</math></b>	4.11	4.11	4.04	3.66	4.64	3.52	4.19	9.23	9.17	15.42
<b>MSE <math>\times 10^4</math></b>	0.14	0.14	1.22	3.88	2.22	51.06	2.72	0.65	1.89	5.49
<b>AARD%</b>	2.53	2.53	1.59	4.17	2.75	3.73	0.45	0.94	1.45	1.70



**Fig 4.** From left to right: plots of lanoline concentration in the gas phase  $C_g$  vs time,  $C_g$  vs bed length ( $z$ ), and extraction rate ( $r_v$ ) vs time. From top to bottom: 120 bar, 80°C, 3 kg/h and 227 kg/m<sup>3</sup>; 120 bar, 80°C, 4 kg/h, 227 kg/m<sup>3</sup>; and 200 bar, 60°C, 4 kg/h and 127 kg/m<sup>3</sup>. Darker lines are those nearer to extractor exit. For middle plot, darker lines are those at latest time.

Table 3 shows the values of MSE. Also, the AARD% are always less than 5%, thereby indicating that the model fits the experimental results very well.

### 6.2 Results on the lanoline solubility in the fluid.

Using eqn. (3), it is possible to calculate the initial solubility of lanoline in the extraction solvent taking for  $C_L$ ,  $C_{L0}$ , the density of lanoline at  $t = 0$ . Table 4 is a summary of our findings about  $C_g^*$  calculated in the present work and the values published by Eychenne et al. (2001) for the same solvent for the same variable settings. As can be seen in the table the differences between the  $C_g^*$  values are of only a few hundredths.

**Table 4. Comparison of lanoline solubility in 5%ethanol-95%CO<sub>2</sub> determined in the present work with that experimentally determined by Eychenne et al. (2001).**

<b>P (bar)</b>	<b>120</b>						<b>150</b>		<b>200</b>	
<b>T (°C)</b>	<b>80</b>						<b>60</b>	<b>80</b>	<b>60</b>	
<b>Q (kg/h)</b>	<b>3</b>		<b>4</b>			<b>5</b>		<b>4</b>		<b>4</b>
<b><math>\rho_B</math> (kg/m<sup>3</sup>)</b>	<b>127</b>	<b>227</b>	<b>159</b>	<b>227</b>	<b>318</b>	<b>127</b>	<b>227</b>	<b>127</b>	<b>127</b>	<b>127</b>
<b><math>C_g^*</math> (kg/m<sup>3</sup>) determined here</b>	0.33	0.33	0.33	0.30	0.38	0.29	0.34	0.76	0.74	1.27
<b><math>C_g^*</math> (kg/m<sup>3</sup>) Eychenne et al.*</b>	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.74	0.69	1.24

\*0.74 and 1.24 are interpolated from all the values given by Eychenne et al.

### 6.3 Results on the fluid-side mass transfer coefficient, $k_G$ .

In order to define the individual mass transfer coefficient, and to check if our values are plausible, the  $k_G$  and the area  $a$  have been separated using the expression for the area given by Valverde and Recasens (2019). This is

$$a = \frac{2(1-\epsilon)}{r} \quad (27)$$

Where  $r$  is the external radius of the cylindrical wool fibre with lanoline, which is actually a variable. According to these authors, the radius  $r$  decreases progressively from an initial value of 10  $\mu\text{m}$  to a final 8.3  $\mu\text{m}$  as extraction takes place (see Table 1). Taking the initial value of  $r$  for calculating an average  $a$ , the values of the individual mass transfer coefficient are given in Table 5.



**Table 5. Fluid-side mass transfer coefficient and mass transfer area calculated from  $k_G$  given in Table 3.**

<b>P (bar)</b>	<b>120</b>						<b>150</b>		<b>200</b>	
<b>T (°C)</b>	<b>80</b>						<b>60</b>	<b>80</b>	<b>60</b>	
<b>Q (kg/h)</b>	<b>3</b>		<b>4</b>			<b>5</b>		<b>4</b>		<b>4</b>
<b><math>\rho_B</math> (kg/m<sup>3</sup>)</b>	<b>127</b>	<b>227</b>	<b>159</b>	<b>227</b>	<b>318</b>	<b>127</b>	<b>227</b>	<b>127</b>	<b>127</b>	<b>127</b>
<b><math>\epsilon</math> (m<sup>3</sup><sub>void</sub>/m<sup>3</sup><sub>bed</sub>)</b>	0.90	0.83	0.88	0.83	0.76	0.90	0.83	0.90	0.90	0.90
<b><math>k_G</math> (x10<sup>6</sup>), m/s</b>	0.32	0.18	8.54	13.72	7.74	12.92	7.80	2.47	2.75	2.27
<b><math>a</math> (x10<sup>-4</sup>), m<sup>2</sup>/m<sup>3</sup></b>	1.93	3.46	2.42	3.46	4.84	1.93	3.46	1.93	1.93	1.93

Puiggené et al. (1997) measured the global mass transfer coefficients,  $K_G$ , in the case of liquid-to-SCF mass transfer and correlated them with the solvent Reynolds number. The values obtained in this work agree within an order of magnitude with those compiled by Puiggené et al. (1997) (see their Fig 12). It is also observed that the mass transfer coefficient  $k_G$  decreases as the extraction pressure increases. This is the expected behaviour for individual mass transfer coefficients (Brunner, 1994). It is also found that the mass transfer coefficient is not greatly affected by temperature. This is the expected feature.

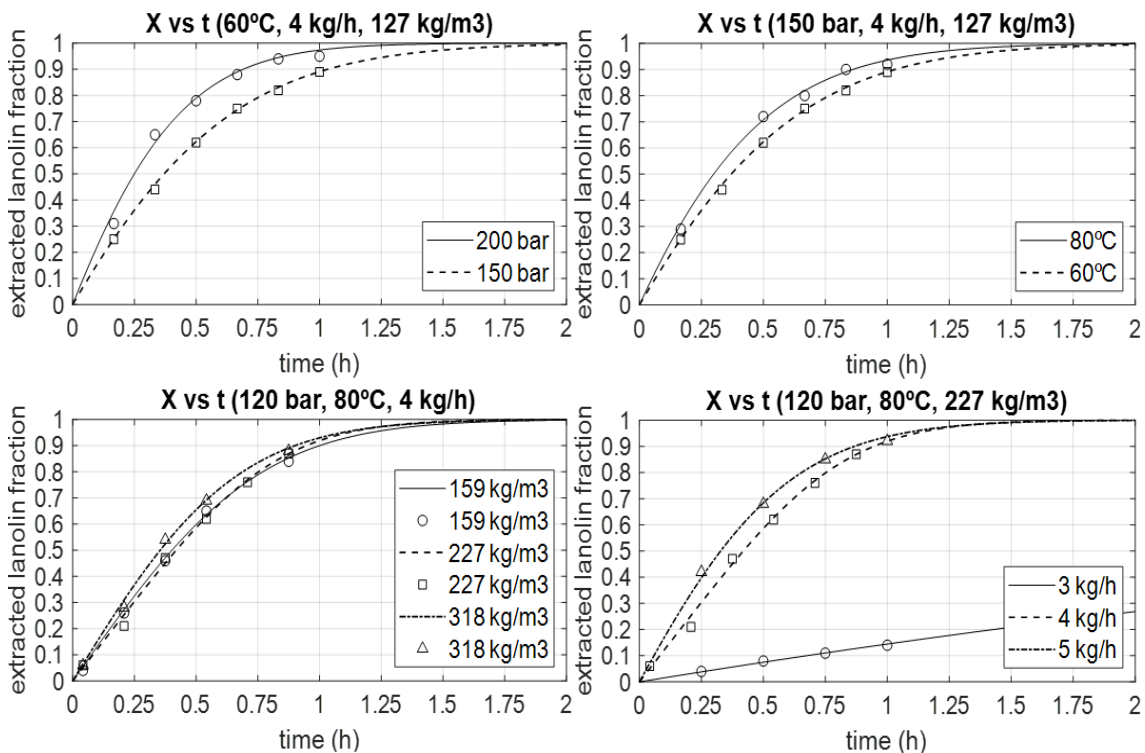
Further comments on the mass transfer coefficient are in order. In a previous work (Valverde and Recasens, 2019), we saw that the Reynolds number in the runs performed by Eychenne et al. (2001) was about  $Re \sim 1$ . Therefore, if we look at the values for  $K_G$  compiled by Puiggené et al. (1997) for  $Re \sim 1$ , the overall mass transfer coefficients are around  $10^{-6}$  and  $10^{-5}$  m/s, hence very close to those reported by King et al. (1987) for the extraction of rapeseed oil with near-critical CO<sub>2</sub>. The values of King et al. agree well with those obtained by Valverde and Recasens (2019) for solid lanoline extraction under similar working conditions (30°C, packing density = 127 kg/m<sup>3</sup> and  $Q = 4$  kg/h),

where values of  $k_G = 5.66 \times 10^{-6}$  m/s at 70 bar and  $k_G = 1.51 \times 10^{-6}$  m/s at 150 bar are found. In this case the inverse effect of pressure, discussed above, is also observed.

As for the effect of packing density, it is seen a slight increment of  $k_G$  in most cases, except for 120 bar, 80°C 4 kg/h and 227 kg/m<sup>3</sup>, where a maximum appears (see the original in Eychenne et al, 2001). The most significant effect is due to the mass flow rate, as noted next. At settings of 3 kg/h there is a markedly lower value than at higher flow rates. However, at 4 and 5 kg/h the difference is less important. As for the effect of temperature this is negligible at increasing flowrates, as expected.

#### 6.4 Performance of the proposed model. Comparison with the neural networks model.

The effects just discussed are seen graphically in the following figures. Observe the fitting quality obtained with *Genetic Algorithm*.



**Fig 5.** Model curves vs experimental points with different state variables. Note that the curve at 120 bar, 80°C, 3 kg/h and 227 kg/m<sup>3</sup> (bottom right plot) is equal to 120 bar, 80°C, 3 kg/h and 127 kg/m<sup>3</sup> because the experimental data are identical.

As seen in the figure it is clear that the fitting of the theoretical curves to the experimental data of extracted fraction vs. time is very good, either for the effects of temperature and pressure or for the effect of packing density and solvent flow rate. For the packing density, however, this is not too evident. Eychenne's original work indicates that the effect of packing density displays quite a strange behaviour that requires further experimental study. The curve with higher fitting error is that at 120 bar, 80°C, 4 kg/h and 127 kg/m<sup>3</sup>. For this curve, for these conditions an MSE = 51 x 10<sup>-4</sup>, is 10 times larger than for the other runs. The reason for this behaviour may be due to experimental error in the extraction runs.

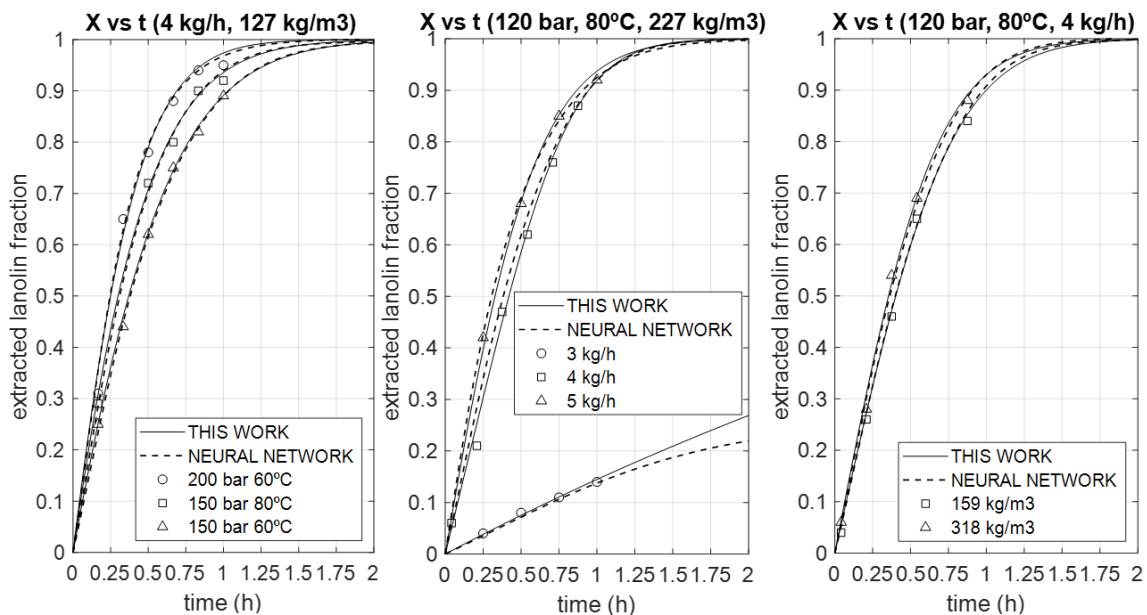
Finally, we have compared the results obtained by Valverde et al. (2019) using neural networks with those of the work done here, in terms of the extracted fraction of lanoline vs time. Recall that in our previous work, the X(t) vs time and the breakthrough curve were based on the use of the integral of the Weibull distribution function (this was reasoned in Valverde et al, 2019). This approach was very useful in modelling not only the breakthrough curve for lanoline, but also for augmenting the rather scarce data of the extraction runs. Note that using the Weibull function, it is possible to calculate a value of K by direct application of eqn. (3), because C<sub>g</sub>\* and C<sub>L0</sub> are known at the start. In Table 6 we compare the values of K obtained in our prior work (Valverde et al, 2019) and those obtained in the present work based on solving the model.

**Table 6. Comparison between Henry's constant, K, obtained in this work and that using the Weibull function (Valverde et al., 2019)**

<b>P (bar)</b>	<b>120</b>						<b>150</b>		<b>200</b>	
<b>T (°C)</b>	<b>80</b>						<b>60</b>	<b>80</b>	<b>60</b>	
<b>Q (kg/h)</b>	<b>3</b>		<b>4</b>			<b>5</b>		<b>4</b>		<b>4</b>
<b>ρ<sub>B</sub> (kg/m<sup>3</sup>)</b>	<b>127</b>	<b>227</b>	<b>159</b>	<b>227</b>	<b>318</b>	<b>127</b>	<b>227</b>	<b>127</b>	<b>127</b>	<b>127</b>
<b>K x 10<sup>4</sup>, (this work)</b>	4.11	4.11	4.04	3.66	4.64	3.52	4.19	9.23	9.17	15.42
<b>K x10<sup>4</sup>, (Weibull method Valverde et al.) (x10<sup>4</sup>)</b>	0.70	0.67	4.12	3.54	4.61	5.06	4.58	8.79	8.42	15.27

The table shows that at 3 kg/h there is an important difference between the K values, most likely because of an insufficient number of experimental data points. For the rest of the cases, however, the agreement is very good. This indicates that the statistical Weibull function used can predict the value of the K constant accurately, without the need to solve the physical model.

In Fig 6, shown next, the extraction curves predicted by the neural network approach, applied by Valverde et al. (2019), are compared with those calculated in the present work. As can be seen, the agreement is excellent.



**Fig 6. Summary of the comparison between the results from neural networks (Valverde et al., 2019) and the curves obtained in this work.**

A final comment on the results of Fig 6 is needed. It is seen that the curves obtained in this work almost coincide with those based on the neural network predictions. The case of 3 kg/h (middle curve) is the only exception. The prediction fails for times of more than 1h. This is the possible cause for the strange value of K, discussed above in connection with Table 6. In general, however, the agreement is very good.

## 6. Conclusions

The solution of the mass balance equations for supercritical extraction, carried out with the finite difference algorithm, provides a method for the design and simulation of the extraction of lanoline from raw wool. The two physical parameters of the model are the Henry-type constant,  $K$ , and the fluid-side mass transfer coefficient,  $k_G$ . After the optimisation of parameters, the model provides a maximum error of 5% with an average error of 2.2 %. We can say that these parameters were not previously available.

The values of the parameters obtained are considered correct, based on the analysis of similar cases found in the literature. Knowledge of the lanoline mass transfer and solubility parameters allows designing from first principles the extraction of lanoline from raw wool with quasi-critical  $\text{CO}_2$ .

The results of the present work firmly validate the methods previously employed by the authors, using neural networks with the Weibull function for an accurate description of the lanoline breakthrough at extractor bed exit.

## Nomenclature

$a$	Mass transfer area, $m^{-1}$
$C_g$	Lanoline conc. in the fluid, $kg/m^3$
$C_g^*$	Lanoline conc. In equil. with $C_L$ , $kg/m^3$
$C_L$	Average Liquid conc., $kg/m^3$
$C_{L0}$	Initial liquid loading, $kg/m^3$
$h_t$	Time step size, s
$h_z$	Space step size, m
$K$	Equilibrium constant, eqn. (3)
$k_g$	Fluid-side mass transfer coefficient, m/s
$P$	Pressure, bar
$Q$	Mass flow rate, $kg/s$
$r_v$	Volumetric mass transfer rate, $kg/m^3 s$
$s$	Laplace variable, $s^{-1}$
$t$	Time, s
$T$	Temperature, $^{\circ}C$
$u$	Superficial velocity, m/s
$V$	Total bed volume, $m^3$
$V_L$	Liquid holdup, $m^3$
$X(t)$	Extracted fraction at time $t$ , -
$z$	Length coordinate, m
$z_T$	Total bed length, m

**Greek symbols and acronyms**

$\varepsilon$	Gas holdup, -
$\varepsilon_L$	Liquid holdup, -
$\rho_B$	Wool packing factor, kg/m <sup>3</sup>
$\tau$	Dummy time variable, s
SCF	Supercritical fluid
SCFE	Supercritical fluid extraction
AARD%	Absolute average relative deviation, eqn. (26)
MSE	Mean squared error, eqn. (24)

## References

Arai, Y., Sako, T., Takebayashi, Y. (Eds.), (2002). Supercritical fluids. Molecular interactions, physical properties and new applications, Springer-Verlag, Berlin.

Bayona, J.M., Erra, P., Moldovan, Z, Domínguez, C., Jover, E, Recasens, F., Larrayoz, M.A., (2005), Method for obtaining lipid fractions from wool or lanoline using pressurized carbon dioxide. Spanish Patent and Trademark Office, Madrid, Patent WO 2002100990A1, 16.07.2005.

Brunner, G., (1994) Gas extraction. An introduction to fundamentals of supercritical fluids and application to separation science, Springer-Verlag, Berlin.

Clifford, A., (1999), Fundamentals of supercritical fluids, Oxford Science Publications, Oxford.

Cygnarowicz-Provost, M., King, J. W., Marmer W. N., Magidman, P., (1994), Extraction of woolgrease with supercritical carbon dioxide, J Amer. Oil Chem. Soc. 71, 223-225.

Domínguez, C., E. Jover, F. Garde, J. Bayona & P. Erra, «Characterization of supercritical fluid extracts from raw wool by TLC-FID and GC-MS, J Amer. Oil Chem. Soc. 80, pp. 717-724, 2003.



Domínguez, C., P. Erra, P., Bayona, J.M., (2010) Physico-chemical and dyeing properties of raw wool extracted by pressurized CO<sub>2</sub>/Modifiers, *Textile research*, 80, 651-659.

Eychenne, V., S. Sáiz, F. Trabelsi & F. Recasens, (2001), Near-critical solvent extraction of wool with modified carbon dioxide - experimental results, *J. Supercrit. Fluids*, 21, 23-31.

Fullana, M., Trabelsi, F., Recasens, F., (1999) Use of neural net computing for statistical and kinetic modelling and simulation of supercritical fluid extractors, *Chem. Eng. Sci.*, 54, 5845-5862.

Fogler, H.S. (2006), *Elements of chemical reaction engineering*, 4th ed., Prentice Hall, Upper Saddle River, NJ, USA

Jones, F., Bateup, D., Dixon, D. and Gray, S., (1995), Solubility of wool wax in supercritical carbon dioxide, *J. Supercrit. Fluids*, 10, 100-111.

King, M.B, and Bott, T.R., (1993), *Extraction of natural products using near-critical solvents*, Chapman & Hall, London, UK.

King, M.B., Bott, T. R., Barr, M. J., Mahmud, R. S., (1987), Equilibrium and rate data for the extraction of Lipids using compressed carbon dioxide, *Sep. Sci. Technol.*, 22 (2), 1103, 1987.

Kuo,B., Kim, J., Jeon,J., Bae,H. (1992), Desorption of wool grease from greasy wool with supercritical carbon dioxide, *Hwahak Konghak*, 30, 491-498.

López-Mesas, M., F. Carrillo, M. Gutiérrez & M. Crespi,(2007). Alternative methods for wool wax extraction from wool scouring wastes, *Aceites y grasas*, 58, 402-407.

McHugh and Krukonis, V., (1996), *Supercritical fluid extraction*, Butterworths, Stoneham MA, USA.

Puiggené, J., Larrayoz, M.A., Recasens, F., (1997), Free liquid-to-supercritical fluid mass transfer in packed beds, *Chem. Eng. Sci.*, 52(2),195-212.

Valverde, A., Recasens, F., (2019), Extraction of solid lanoline from raw wool with near-critical ethanol-modified CO<sub>2</sub>—A mass transfer model, *J. Supercrit. Fluids*,145,151-161.

Valverde, A., Álvarez-Florez, J, Recasens, F., (2019) Application of neural computing to supercritical extraction—Lanoline extraction from raw wool with near-critical CO<sub>2</sub>, *Neural Comp.& Applic.*, under revision.

### 3.5. Assessment of the Conformational Profile of Bombesin by

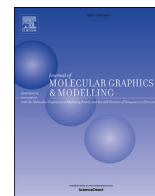
#### Computational Methods

(Published in the Journal of Molecular Graphics and Modelling, in press)

The final contribution of this thesis is aimed at assessing the conformational profile of bombesin, a peptide isolated from European frog *Bombina bombina*, mainly interesting for its applications in drug design, particularly for its role in enzymatic processes related to cancer development. Bombesin shows affinity for the BB2 GPCR, lower affinity for BB1 and non-affinity for the BB3 receptor [125], evidencing a better conformation profile for binding. For this purpose, a 4 $\mu$ s accelerated molecular dynamics (AMD) simulation in water was carried out, reporting in this last paper the results from the Free Energy Landscape (FEL) and the cluster analysis obtained after the Principal Component Analysis (PCA) of the AMD data.

Previous works providing NMR results in 2,2,2-trifluoroethanol/water mixture (30% v/v) by J. A. Carver and J. G. Collins [128] and Díaz et al. [129], and similar Bombesin computational studies by Sharma et al. [130] were referenced to discuss the obtained results. These results agree with previous reported experimental NOE values in the existence of a most populated C-terminus  $\alpha$ -helix structure. However, a less populated but significant N-terminus hairpin structure is obtained from MD results, showing probable fluctuations between different minima basins in the FEL, due to the overlapping of those basins, leading to intermediate conformations observed both in third and fourth clusters results. NOE comparison between experimental and AMD results also shows a significant contribution of side chains of Trp<sup>8</sup> and His<sup>12</sup> in  $\alpha$ -helix folding, taking part on peptide binding whether interacting forces between them are established or not.

This last paper explore the role of statistical mechanics and statistical tools such as PCA to obtain novel results useful for medical purposes, as Bombesin characterization might be. Hence, this work has been published in the *Journal of Molecular Graphics and Modelling*, 2<sup>nd</sup> quartile in Mathematical & Computational Biology area, IF (2018) 1,863, with no citations so far [123].



# Assessment of the conformational profile of bombesin by computational methods

Abel Valverde, Patricia Gomez-Gutierrez, Juan J. Perez\*

Dept. of Chemical Engineering, ETSEIB, Universitat Politècnica de Catalunya, Av. Diagonal, 647, 08028, Barcelona, Spain

## ARTICLE INFO

### Article history:

Received 15 January 2020

Received in revised form

17 March 2020

Accepted 18 March 2020

Available online xxx

### Keywords:

Bombesin

Conformational analysis

Molecular dynamics

Structure-activity analysis

Bioactive peptide

## ABSTRACT

In the present work, the results of a computational study aimed at assessing the conformational profile of bombesin are reported. The conformational space of the peptide was sampled by means of a  $4 \mu\text{s}$  accelerated molecular dynamics simulation in water, using an explicit solvent model. The results were analyzed using Principal Component Analysis to get essential information on peptide fluctuations, along with cluster analysis to characterize different conformations in the sample. Analysis of the results suggests that the peptide adopts helical structures at the C-terminus that tend to unwind at the end of the peptide chain, since there are many structures exhibiting only two turns of a helix at the central segment of the peptide. In addition, the peptide also adopts hairpin turn structures at the N-terminus. Results of the simulation were confronted with available NMR results in a 2,2,2-trifluoroethanol/water (30% v/v) solution. Distances deduced from NOEs experiments only provide support to the presence of helical conformations that represent the most populated structures in the simulation. The absence of other conformations in the NMR experiments can be explained to be due to the  $\alpha$ -helix enhancing nature of the solvent used in the experiments.

© 2020 Elsevier Inc. All rights reserved.

## 1. Introduction

Bombesin, a tetradecapeptide with the sequence:  $\text{Glp}^1\text{-Gln}^2\text{-Arg}^3\text{-Leu}^4\text{-Gly}^5\text{-Asn}^6\text{-Gln}^7\text{-Trp}^8\text{-Ala}^9\text{-Val}^{10}\text{-Gly}^{11}\text{-His}^{12}\text{-Leu}^{13}\text{-Met}^{14}\text{-NH}_2$  (Glp = pyroglutamic acid) was originally isolated from the skin of the European frog *Bombina orientalis* [1]. After its discovery, other peptides with high sequence identity like Ranatensin, Alytensin, Phyllolitorin or Litorin among others, were subsequently characterized [2,3]. Members of this family can be classified into three groups according to the sequence of the C-terminus: a first group is comprised by the bombesin related peptides with sequence:  $\text{Gly-His-Leu-Met-NH}_2$ ; a second group is comprised of the litorin-ranatensin related peptides with sequence  $\text{Gly-His-Phe-Met-NH}_2$ ; and a third group comprising the phyllolitorin related peptides with sequence:  $\text{Gly-Ser-Phe(Leu)-Met-NH}_2$ . Despite these peptides were originally isolated from the skin of diverse amphibians, it was later found that they are also widely distributed in mammals [4]. Specifically, only two peptides of this family have been isolated in mammals until now: Neuromedin B (NMB) [5], a

member of the litorin-ranatensin family and the Gastrin-releasing peptide (GRP) [6], a 27 residue long peptide together with its short version retaining full activity, the GPR(18–27) -also known as Neuromedin C (NMC)- that belongs to the bombesin related peptides group. These two peptides have homologous C-terminal segments and share an identical C-terminal heptapeptide.

These peptides are involved in a wide spectrum of biological activities both, in the central nervous system including satiety, control of circadian rhythm, thermoregulation and in peripheral tissues, including stimulation of gastrointestinal hormone release, activation of macrophages and effects on development [7]. Actions of these peptides are mediated through three GPCRs:  $\text{BB}_1$ ,  $\text{BB}_2$  and  $\text{BB}_3$  [8]. Bombesin exhibits nanomolar affinity for the  $\text{BB}_2$  receptor and about ten times lower affinity for  $\text{BB}_1$ , showing no affinity for the  $\text{BB}_3$  receptor. In contrast, NMB exhibits nanomolar affinity for  $\text{BB}_1$  and about fifty times lower affinity for  $\text{BB}_2$  receptor [9]. The  $\text{BB}_3$  is considered an orphan receptor, since no endogenous ligand has yet been identified. It is involved in energy balance, glucose homeostasis, control of body weight or tumor growth. Due to the wide spectrum of biological activities mediated by the bombesin receptors, there is a considerable interest in the clinical potential of novel agonist and/or antagonist, particularly in the fight against cancer [10,11]. Also, due to its role in the control of appetite,

\* Corresponding author.

E-mail address: [juan.jesus.perez@upc.edu](mailto:juan.jesus.perez@upc.edu) (J.J. Perez).

metabolism, and chronic itching they are interesting targets for drug discovery [12,13]. In order to develop new drugs targeting bombesin receptors, it is necessary a deeper understanding of the structure-activity relationships of these peptides. Specifically, identification of the features of the bioactive conformation is necessary to understand the relative position of key residues involved in recognition and activation [14,15].

In regard to bombesin, it was soon demonstrated that the analog bombesin(6–14) is the shortest peptide sequence retaining full activity [16,17]. Moreover, the synthesis and biological evaluation of diverse peptide analogs soon demonstrated that residues Trp<sup>8</sup> and His<sup>12</sup> are important for the activity of the peptide [18] and that Met<sup>14</sup> is key for activity, since its deletion transforms the peptide into a potent antagonist [19]. In addition to identify key residues necessary for the biological activity of bombesin, it is also important to understand the conformational features of the peptide. Concerning its structure, spectroscopy studies including NMR [20–24], Infrared (IR) [25], Circular Dichroism (CD) and Fluorescence spectroscopy [26] with the use of different solvents like water, dimethylsulfoxide (DMSO) or 2,2,2-trifluoroethanol-water mixtures have provided information of the conformational features of bombesin. Specifically, NMR reports in water and DMSO [20–22] describe the structure of bombesin as a random coil. In contrast, NMR experiments of bombesin carried out in a 2,2,2-trifluoroethanol/water mixture (30% v/v) [23,24] report that the C-terminal segment of the peptide ranging from residue 6 to 14 displays a helical conformation, although with residues 11 to 14 less sharply structured. Moreover, it was also concluded that the first two N-terminal residues adopt an extended conformation, while the region between residues 3 and 5 exhibits a great deal of flexibility. IR [25], CD and Fluorescence studies [26] also confirm the adoption of a helical conformation when the peptide is incorporated into lipid environments.

Computational studies addressed to assess the conformational profile of bombesin have been also reported in the past [27]. Thus, Replica Exchange Molecular Dynamics simulations in implicit solvent showed that the peptide attains a helical structure on the segment bombesin(6–14) with a tendency to unwind at the C-terminus. Moreover, these results also show a frequent occurrence of conformations that bring together the side chains of aromatic residues Trp<sup>8</sup> and His<sup>12</sup> in close proximity, fact that was considered important to explain the importance of these two residues in peptide binding. However, there still open questions regarding the conformational profile of bombesin like, does the peptide adopt any other significant conformation in addition to the helical one? Or, do residues Trp<sup>8</sup> and His<sup>12</sup> play a role in the stability of the helical structure? Access to faster computers, as well as to more efficient algorithms permits nowadays to perform thorough samplings of flexible molecules. In the present study, we report the results of a detailed computational study of bombesin in water, using accelerated molecular dynamics (aMD) as sampling technique in explicit solvent. Present results confirm the tendency of the peptide to adopt a helical conformation in the middle and C-terminal segments. In addition, present results also show a less populated conformation where the peptide adopts a hairpin turn structure at the N-terminus of the peptide.

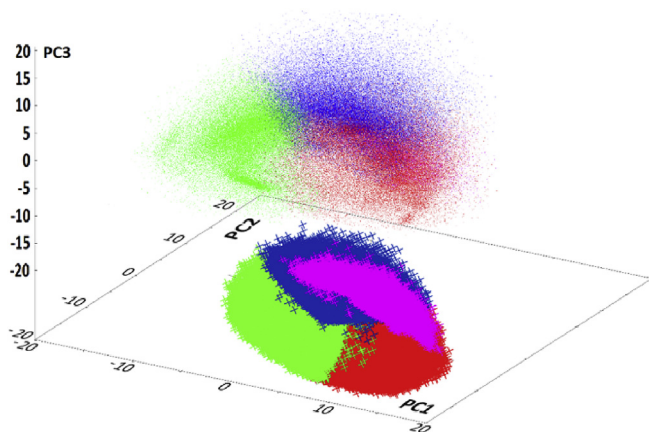
## 2. Methods

The starting structure of the peptide in its extended conformation was generated by means of the LEaP program embedded in the Amber 12 software [28]. Subsequently, the structure was soaked in a rectangular cuboid box of equilibrated TIP3P water molecules [29]. The system, composed by the peptide and 25140 water molecules, was neutralized by the addition of one chloride

ion. Conformational sampling was carried out after a 4  $\mu$ s aMD simulation at 300 K, following the protocol previously proposed by McCammon et al. and implemented in Amber 12 [30,31]. Specifically, a biased potential is applied to the total energy and the dihedral torsional energy. The simulation was carried out in the NVT collective using the ff99SB force field [32] at 300K and periodic boundary conditions. A cutoff of 10 Å was used for the calculation of the non-covalent interactions and electrostatic interactions were treated using the PME method. Before starting the aMD simulation, the structure was energy minimized through 2500 steps of the steepest descent method followed by a few iterations using the conjugate gradient algorithm in order to remove possible steric clashes. In a subsequent step the system was heated up to 300K using a 100 ps NVT molecular dynamics (MD) calculation at a rate of 30 K per 10 ps and equilibrated through 1 ns of NPT MD and 1 ns of NVT MD. At this point a 10 ns NVT MD was performed in order to determine the average potential energy ( $E_p$ ), the mean dihedral energy ( $E_d$ ), as well as the parameters  $\alpha_P$  and  $\alpha_D$ , necessary for the subsequent aMD simulation. During the simulations, the temperature was kept at 300 K through the Langevin thermostat with a collision frequency of 2 ps<sup>-1</sup>. Moreover, the SHAKE algorithm was used in order to constrain all bonds involving hydrogen atoms allowing us to use an integration step of 2 fs. After completion of the aMD calculation, effects due to the energy bias were removed using a ten order Maclaurin series reweighting for each configuration in order to recover a canonical ensemble [30].

## 3. Results and discussion

The conformational profile of the peptide was assessed using 400,000 structures extracted at regular intervals from the 4  $\mu$ s aMD trajectory. In order to extract the most relevant information about the flexibility of the peptide, the sample was subjected to Principal Component Analysis (PCA) [35] using the cpptraj module embedded in Amber 12 [28]. For this purpose, the 400,000 structures were first superimposed to a structure of reference using the backbone coordinates of the diverse residues without including the N- and the C-terminal ones. Then, the displacements of the backbone atoms in regard to the average structure were used to compute a covariance matrix that was subsequently diagonalized to obtain the set of principal components (PCs) together with their eigenvalues. Once the PCs were rank ordered according to their eigenvalues, the results showed that the first PC captures 23.2% of the variance of the system; the first two PCs capture 40.8% of the variance and the addition of a third component PC3, capture a variance of 52.8%. The low variance captured by PC1 can be attributed to the large flexibility of the peptide, as found in similar studies [36]. Diverse analysis using different number of snapshots were carried out in the present work to rule out a possible bias in the procedure. In a subsequent step, the 400,000 structures were subjected to cluster analysis using of the average linkage algorithm [37]. For this purpose, the backbone C $\alpha$  root-mean-square deviation of residues 2–13 was used as a measure for the distance between two conformations. Comparison of the results obtained using a different number of clusters suggested that four clusters was adequate to explain qualitatively the conformational profile of the peptide. Fig. 1 depicts pictorially the 400,000 structures in the space defined by the three first components PC1– PC3 and its projection onto the (PC1, PC2) plane with the diverse clusters depicted in different colors. Analysis of the Figure permits to observe that although clusters #1, #3 and #4 are overlapped in the (PC1, PC2) map, they can be partially segregated when three PCs are used. Moreover, even using three PCs there is still some overlap between clusters #1 and #4 due to the similarity of the structures contained in both clusters, as described later.

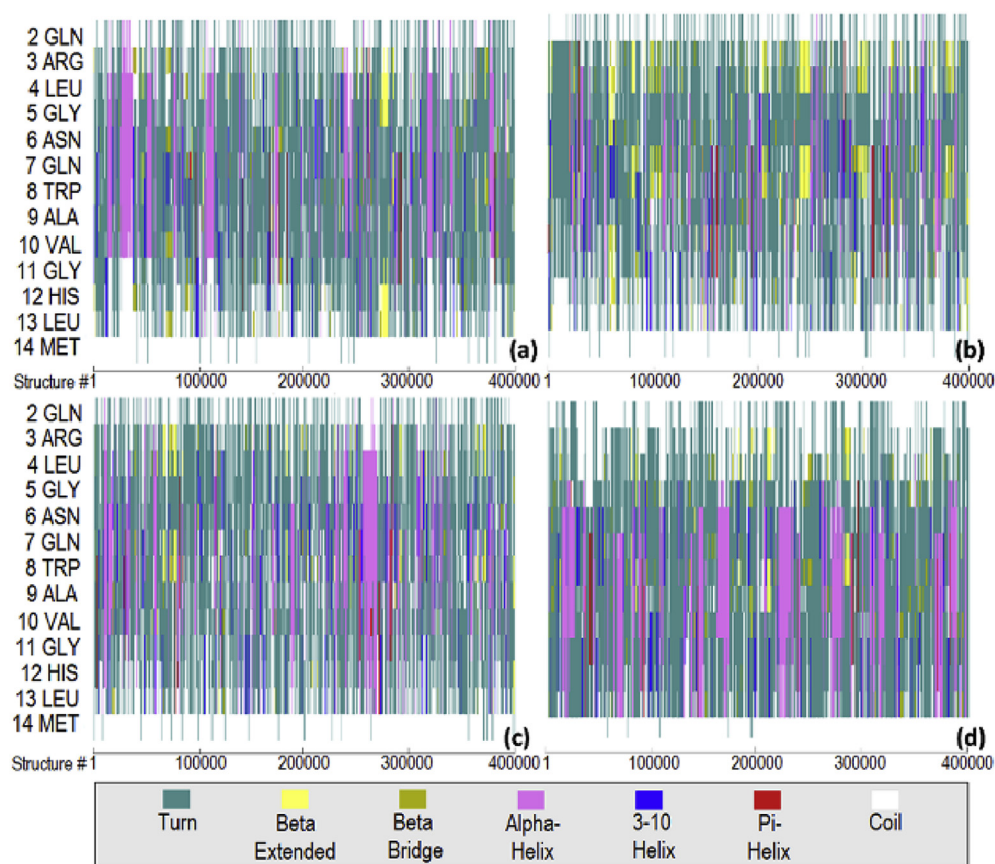


**Fig. 1.** Projection of the 400,000 structures selected for the analysis onto their Principal Components. a) Projection onto the three Principal Components PC1-PC3; b) Projection onto components PC1 and PC2. Structures are depicted in colors according to the cluster they belong: cluster#1 red; cluster#2 green; cluster#3 blue; cluster#4 magenta.

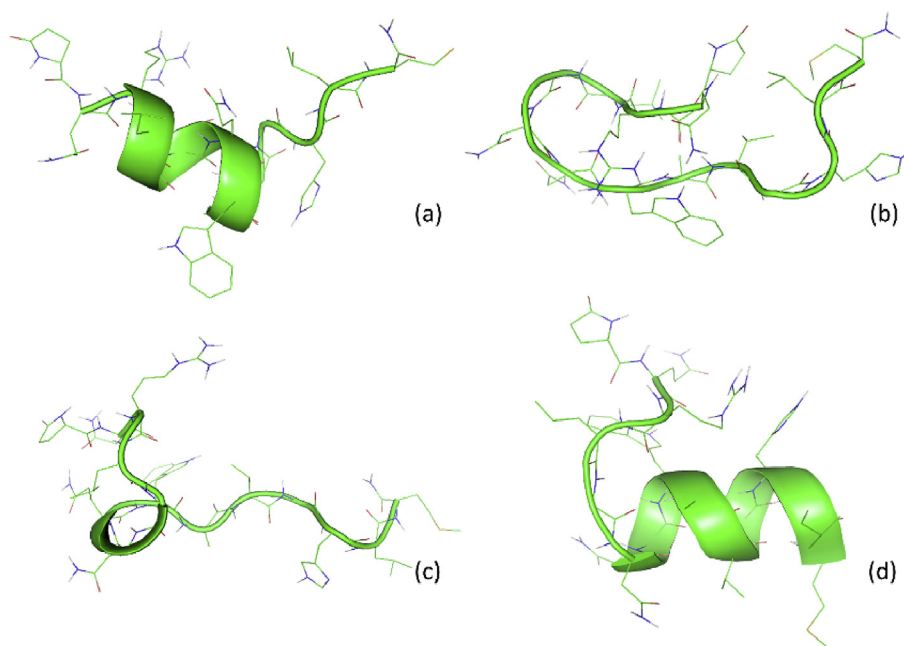
Analysis of the secondary structure contents of the diverse structures of the different clusters permits to obtain a simplified picture of the conformational features of the structures in each cluster. Fig. 2a–d shows for each of the clusters the secondary structure adopted by the diverse residues (Y-axis) in each of the

structures (X-axis), using a color-coded diagram. Thus, analysis of Fig. 2a suggests -considering together helical secondary motifs and turns- that cluster #1 (in red in Fig. 1) includes partially helical structures of diverse lengths at the central segment of the peptide. Fig. 3a depicts pictorially a representative structure of this cluster (the pdb file of the structure is provided as supplementary material). Analysis of Fig. 2b suggests that cluster#2 (in green in Fig. 1) contains structures with a hairpin turn conformation at the N-terminus involving residues Gln<sup>2</sup>-Arg<sup>3</sup>-Leu<sup>4</sup>-Gly<sup>5</sup>-Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-Ala<sup>9</sup>-Val<sup>10</sup>. Fig. 3b depicts pictorially a representative structure of this cluster (the pdb file of the structure is provided as supplementary material). Analysis of Fig. 2c suggest that cluster#3 (in navy blue in Fig. 1) contains intermediate structures that exhibit diverse combination of turns, as for example the structure shown in Fig. 3c (the pdb file of the structure is provided as supplementary material). Finally, analysis of Fig. 2d suggest that cluster#4 (in magenta in Fig. 1) contains helical structures involving residues 6–14, as shown in Fig. 3d that are closely related with the structures of cluster#1 (the pdb file of the structure depicted in Fig. 3d is provided as supplementary material).

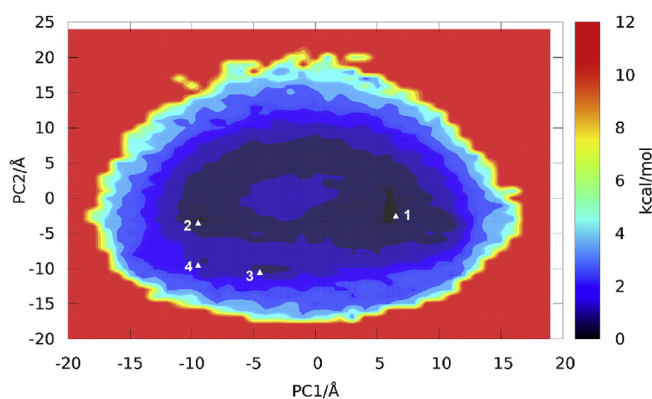
Another interesting result from the representation of the structures in the PCs space is that the density of points provides insight into the preferred conformations visited during the sampling process. Moreover, these points can be associated with the features of the free energy landscape (FEL) of the peptide and consequently, permits to identify low free energy conformations as well as the difference between them [35]. Fig. 4 shows the projection of the 400,000 structures onto the PC1 and PC2 using a color



**Fig. 2.** Secondary structure adopted by the diverse residues (Y-axis) in the 400,000 structures (X-axis) used for analysis. The color code used is: helix (magenta); turn (blueish-green); extended (yellow). a) Structures of cluster#1; b) Structures of cluster#2; c) Structures of cluster#3; d) Structures of cluster#4.

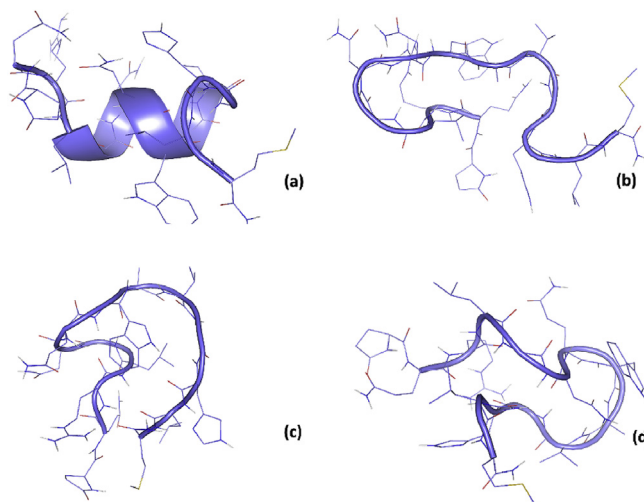


**Fig. 3.** Structures representing the diverse clusters. (a) Cluster#1; (b) cluster#2; (c) cluster#3; (d) cluster#4.



**Fig. 4.** Contour plot of the Free energy landscape of the peptide. Low regions are plotted in dark blue, whereas high regions in red. The location of the four minima is also indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

code that depicts the density of structures contained in a small square used as probe to analyze the space. Inspection of Fig. 4 permits to identify four minima. Minimum #1 is located on cluster#1 at coordinates (6,-2.5) and is the most populated, being representative structure is a partial helix located on the central segment of the peptide similar to the representative structure of cluster#1, as shown in Fig. 5a (pdb file is provided as supplementary material). Minimum #2 is located in the region of cluster#2 at coordinates (-9.5,-3), being its representative structure a hairpin turn at the N-terminus involving residues Gln<sup>2</sup>-Arg<sup>3</sup>-Leu<sup>4</sup>-Gly<sup>5</sup>-Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-Ala<sup>9</sup>-Val<sup>10</sup> similar to the representative structure of cluster#2, as shown in Fig. 5b (pdb file is provided as supplementary material). The representative structure of minimum #3 located at coordinates (-4.5,-10) can be considered an intermediate structure exhibiting a turn of a helix, as shown in Fig. 5c (pdb file is provided as supplementary material). Finally, the representative structure of minimum #4 located at coordinates (-9.5,-9.5) can also be considered as an intermediate structure with diverse turns,



**Fig. 5.** Structures representing the diverse minima identified in the sampling process. (a) Minimum M1; (b) minimum M2; (c) minimum M3; (d) minimum M4.

as shown in Fig. 5d (pdb file is provided as supplementary material). These latter two could be considered as kinetic traps.

Analysis of these results provide an emerging qualitative picture of the conformational profile of bombesin: the peptide attains diverse conformations being the most populated the  $\alpha$ -helix involving diverse residues at the central and C-terminus segments. On the one hand, cluster#4 includes a large population of helical structures involving residues Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-Ala<sup>9</sup>-Val<sup>10</sup>-Gly<sup>11</sup>-His<sup>12</sup>-Leu<sup>13</sup>-Met<sup>14</sup> and on the other, there is the large population of partially helical structures included in cluster #1 that exhibit a few turns of a helix, suggesting the relevance of the helical conformation in the conformational profile of the peptide. These results taken together permit to interpret that the peptide is likely to attend helical structures with a tendency to unwind at the C-terminus. These structures correspond to those identified in a previous work of this laboratory [27] and agrees well with the results



found in NMR experiments [20]. Moreover, the peptide also adopts hairpin turn conformations involving residues at the N-terminus. The rest of the structures correspond to partially folded structures, some of them very populated.

The presence of helical structures at the C-terminus of the peptide could be associated to the activity of the peptide since bombesin(6–14) is the shortest segment retaining full activity [16,17]. However, there are important structural differences between a  $\alpha$ -helix involving residues 6–14 and a partial helix involving a few residues in the middle segment that can be relevant to explain the biological activity of the peptide. Structure-activity studies showed the importance of residues Trp<sup>8</sup> and His<sup>12</sup> for the activity of the peptide. This can be interpreted in structural terms that these residues play a role in stabilizing the bioactive conformation of the peptide or alternatively that they are directly involved in the recognition by its receptors. The distribution of the distance between the centers of Trp<sup>8</sup> and His<sup>12</sup> side chain aromatic rings monitored along the MD trajectory displayed in Fig. 6 shows sample a wide set of distances with a maximum between 8 and 13 Å. Actually, conformations belonging to cluster#2, cluster#3 and cluster#4 exhibit distances between aromatic rings in this range. In contrast, structures belonging to cluster#1 that are partially helical (Fig. 3a) exhibit distances  $\leq 7$  Å. These type of structures permit the side chains of Trp<sup>8</sup> and His<sup>12</sup> to be involved in a quadrupole-quadrupole interaction stabilizing the conformation [38]. Actually, side chain-side chain interactions compensate the loss of helicity at the C-terminus in this type of structures explaining why this conformation is the most populated structure in the sampling, representing minimum #1 on the FEL. The importance of adoption of a partial helix at the middle segment for receptor recognition is further reinforced from the results of activity exhibited by diverse bombesin analogs that incorporate turn mimetics in substitution of residues Val<sup>10</sup>-Gly<sup>11</sup> [39].

The peptide also adopts hairpin turn conformations at the N-terminus involving residues Gln<sup>2</sup>-Arg<sup>3</sup>-Leu<sup>4</sup>-Gly<sup>5</sup>-Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-

Ala<sup>9</sup>-Val<sup>10</sup> (Fig. 3b). Actually, the structure is stabilized by hydrogen bonds between the amide hydrogen of Leu<sup>4</sup> and the carbonyl oxygen of Gln<sup>7</sup> adopting a reverse type II  $\beta$ -turn and between the carbonyl oxygen of Gln<sup>2</sup> and the amide hydrogen of Val<sup>10</sup> that stabilizes the hairpin. This structure corresponds to minimum M2 and it is quite abundant in cluster#2. In these structures the side chains of Trp<sup>8</sup> and His<sup>12</sup> do not interact since distances are longer than 10 Å, although we cannot rule out that may be relevant for receptor recognition. An interesting feature of these structures is that one face of the structure aligns side chains of hydrophobic residues, whereas the other face involves polar side chains. This structure is very likely to be adopted by the peptide in lipid-aqueous interfaces. In the FEL there are other two minima represent intermediate structures and can be considered as kinetic traps.

Results of the sampling can be contrasted with the results derived from NMR studies in 2,2,2-trifluoroethanol/water mixture (30% v/v) [23,24]. Specifically, the intensity of the signals obtained from Nuclear Overhauser Effect (NOE) experiments, which stem from cross-relaxation between proton pairs in NMR experiments are related to the distance between the protons involved, permitting a direct comparison with the results of the molecular dynamics trajectory [33]. According to the literature [34], strong NOE peaks correspond to proton-proton distances in the interval 1,8–2,7 Å; medium NOE peaks in the interval 1,8–3,3 Å, weak NOE peaks in the interval 1,8–5,0 Å and very weak in the interval 1,8–6,0 Å.

Fig. 7 shows in orange the distance intervals derived from NOEs experiments according to their intensity, along with the distance intervals sampled by the peptide during the simulation. Inspection of Fig. 7 shows that all distance intervals derived from NMR experiments are sampled during the simulation, although during the sampling process the peptide also explores alternative distances that are not detected in the experiments. In order to characterize the structures compatible with the distances derived from NMR experiments, Fig. 8a–d shows the same diagram of Fig. 6 depicting distances from the set of structures embedded in

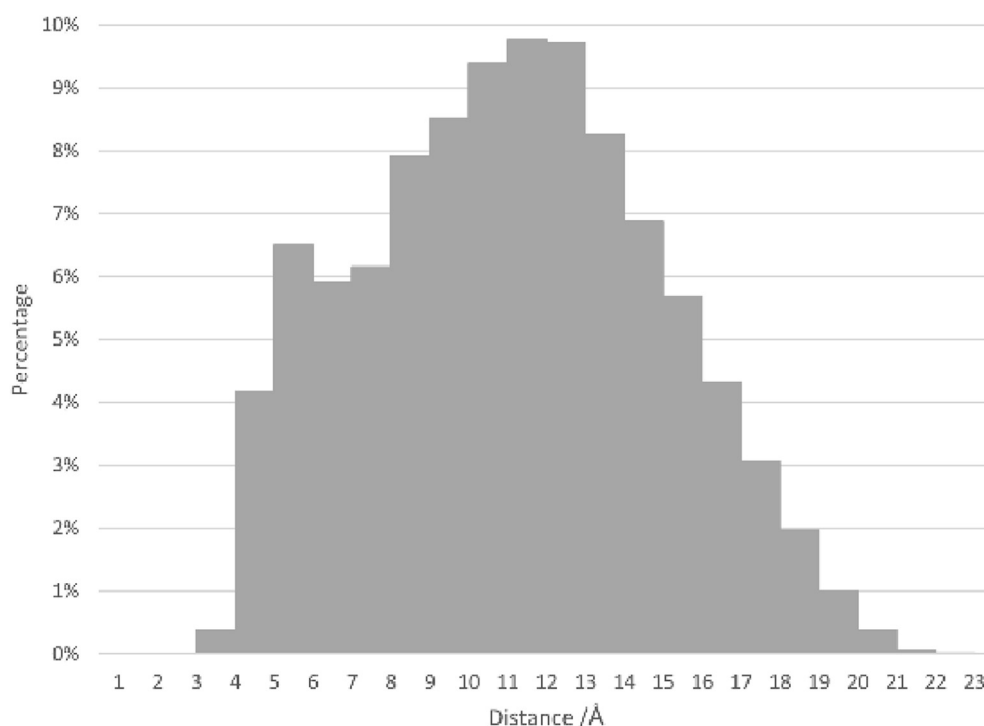


Fig. 6. Distribution of the distance between the centers of the aromatic rings of the side chains of residues Trp<sup>8</sup> and His<sup>12</sup>.

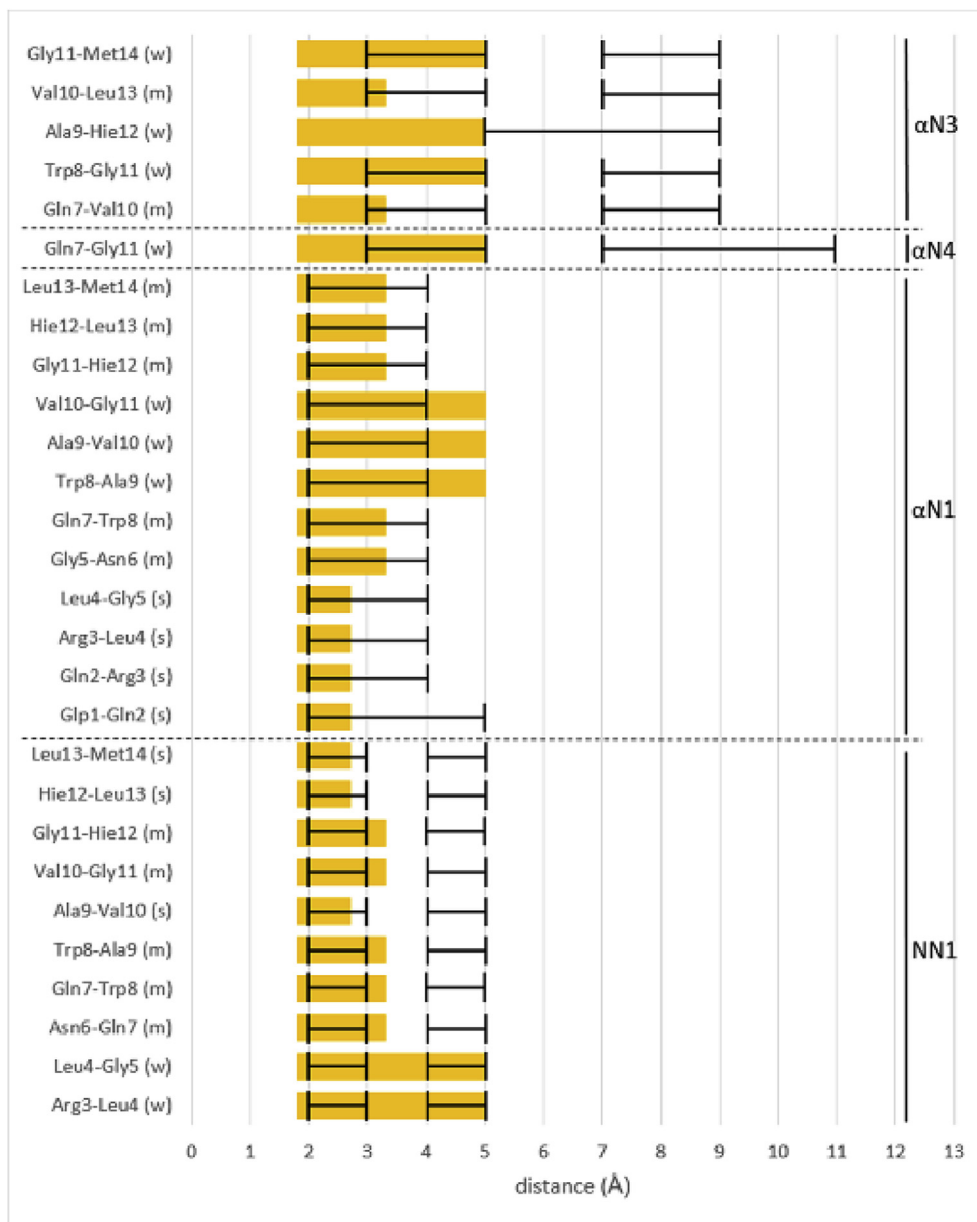


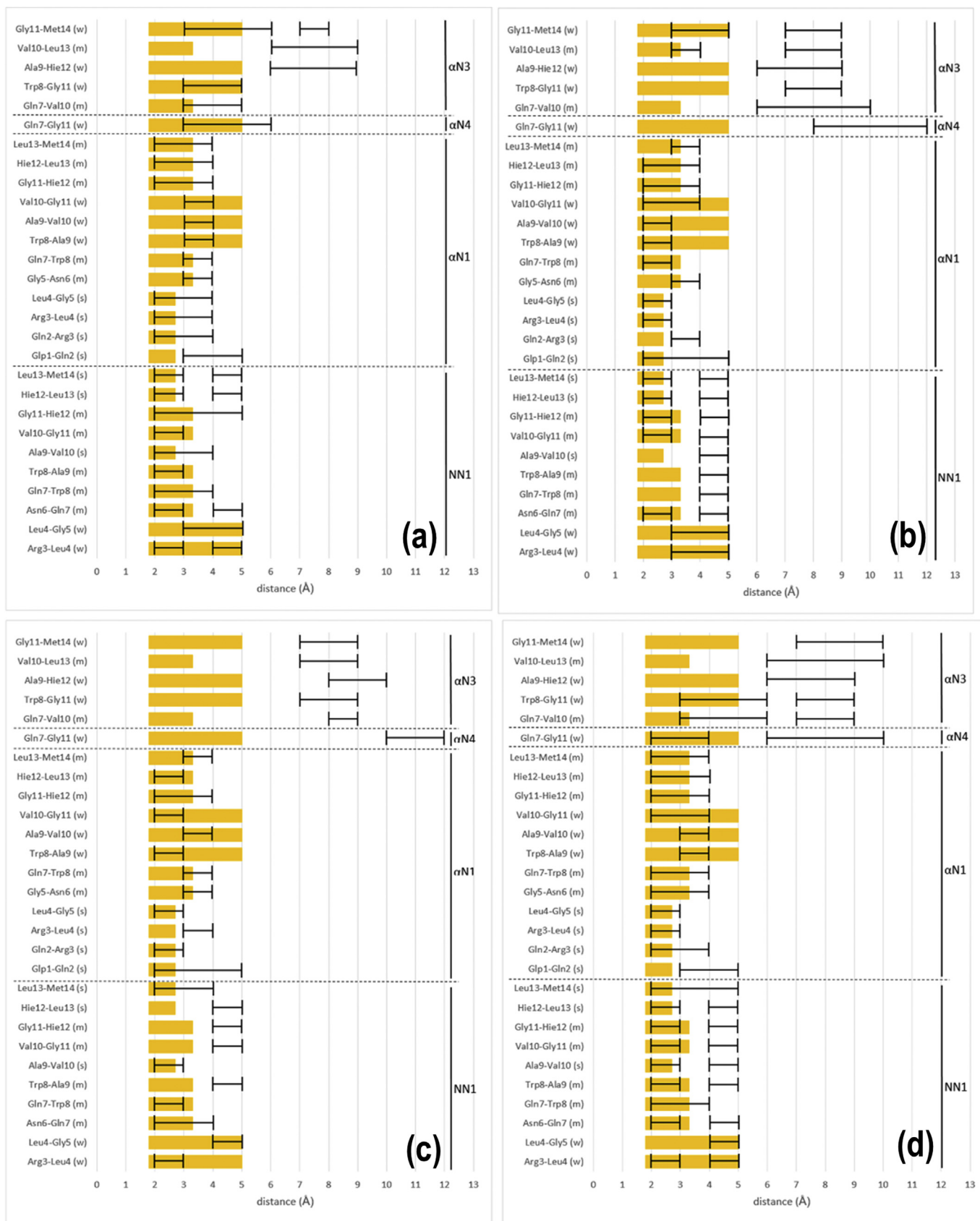
Fig. 7. Comparison of the distances deduced from NOEs experiments of bombesin in trifluoroethanol/water 30% (v/v) (refs. 23 and 24) and sampled during the whole trajectory of the simulation.

the square defining each of the minima in the FEL. As can be seen, structures representing minimum M1 (Fig. 8a) fit better, although none of the structures completely fulfill the distances derived from NOEs experiments. In order to understand better the features of the experimental structure, we measured the corresponding distances on the representative structures of the different clusters (Fig. 3a–d). Interestingly, the representative structure of cluster#4 (Fig. 3d) fits the best, except for the N-terminus residues, clearly indicating that in the structure derived from NMR all residues are in a helical conformation. In order to conceal the results from the simulation and from NMR experiments, it is necessary to consider that a peptide in solution is flexible enough to adopt diverse conformations, although in this case, the  $\alpha$ -helix is a very populated conformation. On the other hand, the helix enhancing nature of the 2,2,2-trifluoroethanol may bias the experiments, inducing more structured

conformations and consequently, increasing the population of helical conformations determined experimentally [40].

From a structure-activity perspective, since the segment bombesin(6–14) is known to retain full activity, helical structures located in the middle and C-terminal segments should be considered relevant for peptide recognition. Moreover, experimental results indicate that Met<sup>14</sup> is key for receptor activation and that Trp<sup>8</sup> and His<sup>12</sup> are relevant for receptor recognition. These results are compatible with a binding mode of bombesin into its receptor through the C-terminus, in such a way that the peptide adopts a helical structure. This permits residue Met<sup>14</sup> to be the most deeply inserted in connection with its role to be involved in activation [41].

In order to understand the type of helical structure adopted by the peptide when bound to its receptor, it is important to analyze the geometries sampled by the side chains of Trp<sup>8</sup> and His<sup>12</sup>. Distances between the centers of the two aromatic moieties of the



**Fig. 8.** Same as Fig. 6 for the distances representing the diverse minima, considering all the structures in the probe square used to draw the FEL of the peptide (Fig. 4). (a) Minimum M1; (b) minimum M2; (c) minimum M3; (d) minimum M4.

respective side chains range from 5 Å to more than 13 Å suggesting that the peptide is quite flexible. Distances more frequently sampled correspond to helical structures and hairpin turns, whereas lower distances correspond to structures of cluster#1 that exhibit a partial helical structure in the middle segment. In these latter structures, the side chains of Trp<sup>8</sup> and His<sup>12</sup> are involved in a quadrupole-quadrupole interactions that permit to compensate the loss of helicity of the C-terminus segment explaining the large population sampled. Accordingly, there are two possible binding modes. In one, the peptide adopts a partial helical structure stabilized by the side chains of Trp<sup>8</sup> and His<sup>12</sup> that permits the side chains of the rest of the C-terminus residues to adopt a conformation suitable for binding. In the other, the peptide adopts a helix, where the side chains of Trp<sup>8</sup> and His<sup>12</sup> do not interact, but adopt a distance suitable for binding to specific residues of the receptor. Unfortunately, the information available does not permit to discard any of the two hypothesis. Further investigations involving modeling of the bombesin receptors are necessary to rule out one of the hypothesis.

#### 4. Conclusions

The conformational profile of bombesin was assessed in the present work through a computational study. Specifically, a 4 μs accelerated molecular dynamics trajectory was used to sample the conformational space. Analysis of the trajectory was carried out by means of Principal Component Analysis that permitted to extract the essential information of the fluctuations of the peptide complemented by cluster analysis that permitted to classify the diverse conformations sampled in different subsets.

The analysis of the sampling reveals the tendency of the peptide to adopt helical structures at the C-terminus of the peptide as the most populated conformation. More specifically, the peptide adopts a  $\alpha$ -helix conformation involving residues Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-Ala<sup>9</sup>-Val<sup>10</sup>-Gly<sup>11</sup>-His<sup>12</sup>-Leu<sup>13</sup>-Met<sup>14</sup> as well as partially helical structures involving two or three helix turns at the central segment of the peptide, being these latter more abundant. The coexistence of these two type of helical conformations can be interpreted considering that the peptide exhibits a  $\alpha$ -helix at the C-terminus that tends to unwind in its last residues. In addition, the peptide adopts hairpin structures at the N-terminus involving residues Gln<sup>2</sup>-Arg<sup>3</sup>-Leu<sup>4</sup>-Gly<sup>5</sup>-Asn<sup>6</sup>-Gln<sup>7</sup>-Trp<sup>8</sup>-Ala<sup>9</sup>-Val<sup>10</sup>. This structure permits to obtain a structure with one side hydrophilic and the opposite hydrophobic. These two structures appear as minima in the FEL of the peptide in this work. The rest of structures can be considered as intermediate structures. Interestingly, there are two of the structures that appear very populated in the FEL and can be considered as kinetic traps. These results agree well with the conclusions of previous NMR experiments carried out in 2,2,2-trifluoroethanol.

The features of the bombesin structure sketched in this work permits to speculate that the peptide binds into its receptor through the C-terminus, in such a way that the peptide adopts a helical structure. This permits residue Met<sup>14</sup> to be the most deeply inserted in connection with its role to be involved in activation.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmgm.2020.107590>.

#### References

- [1] A. Anastasi, V. Erspamer, M. Bucci, Isolation and structure of bombesin and alytesin, 2 analogous active peptides from skin of European amphibians *bombina* and *alytes*, *Experientia* 27 (1971) 166–167.
- [2] V. Erspamer, P. Melchiorri, Active polypeptides - from amphibian skin to gastrointestinal-tract and brain of mammals, *Trends Pharmacol. Sci.* 1 (1980) 391–395.
- [3] R.T. Jensen, T.W. Moody, Bombesin peptides (cancer), in: A.J. Kastin (Ed.), *Hand-book of Biologically Active Peptides*, Elsevier, Amsterdam, 2013, pp. 506–511.
- [4] H. Ohki-Hamazaki, M. Iwabuchi, F. Maekawa, Development and function of bombesin-like peptides and their receptors, *Int. J. Dev. Biol.* 49 (2005) 293–300.
- [5] N. Minamida, K. Kangawa, H. Matsuo, Neuromedin B: a novel bombesin-like peptide identified in porcine spinal cord, *Biochem. Biophys. Res. Commun.* 114 (1983) 541–548.
- [6] T.J. McDonald, H. Jornvall, G. Nilsson, M. Vagne, M. Ghatei, S.R. Bloom, V. Mutt, Characterization of a gastrin-releasing peptide from porcine non-antral gastric tissue, *Biochem. Biophys. Res. Commun.* 90 (1979) 227–233.
- [7] H.C. Weber, Regulation and signaling of human bombesin receptors and their biological effects, *Curr. Opin. Endocrinol. Diabetes Obes.* 16 (2009) 66–71.
- [8] R.T. Jensen, J.F. Battey, E.R. Spindel, R.V. Beny, International union of pharmacology. LXVIII. Mammalian bombesin receptors: nomenclature, distribution, pharmacology, signaling, and functions in normal and disease states, *Pharmacol. Rev.* 60 (2008) 1–42.
- [9] R.V. Benya, T. Kusui, T.K. Pradhan, J.F. Battey, R.T. Jensen, Expression and characterization of cloned human bombesin receptors, *Mol. Pharmacol.* 47 (1995) 10–20.
- [10] S. Mahmoud, J. Staley, J. Taylor, A. Bogden, J.-P. Moreau, D. Coy, I. Avis, F. Cuttitta, J.L. Mulshine, T.W. Moody, [Psi13,14] Bombesin analogs inhibit growth of small-cell lung-cancer in vitro and in vivo, *Canc. Res.* 51 (1991) 1798–1802.
- [11] Y. Qin, T. Ertl, R.-Z. Cai, G. Halmos, A.V. Schally, Inhibitory effect of bombesin receptor antagonist rc-3095 on the growth of human pancreatic-cancer cells in-vivo and in-vitro, *Canc. Res.* 54 (1994) 1035–1041.
- [12] Z. Merali, J. McIntosh, H. Anisman, Role of bombesin-related peptides in the control of food intake, *Neuropeptides* 33 (1999) 376–386.
- [13] K. Yamada, E. Wada, Y. Santo-Yamada, K. Wada, Bombesin and its family of peptides: prospects for the treatment of obesity, *Eur. J. Pharmacol.* 440 (2002) 281–290.
- [14] J.J. Perez, F. Corcho, O. Llorens, Molecular modeling in the design of peptidomimetics and peptide surrogates, *Curr. Med. Chem.* 9 (2002) 2209–2229.
- [15] J.J. Perez, Designing peptidomimetics, *Curr. Top. Med. Chem.* 18 (2018) 566–590.
- [16] S.E. Gargosky, J.C. Wallace, F.M. Upton, F.J. Ballard, C-Terminal bombesin sequence requirements for binding and effects on protein synthesis in Swiss 3T3 cells, *Biochem. J.* 247 (1987) 427–432.
- [17] J.T. Lin, D.H. Coy, S.A. Mantey, R.T. Jensen, Comparison of the peptide structural requirements for high affinity interaction with bombesin receptors, *Eur. J. Pharmacol.* 294 (1995) 55–69.
- [18] M. Broccardo, G. Falconieri Erspamer, P. Melchiorri, L. Negri, R. De Castiglione, Relative potency of bombesin-like peptides, *Br. J. Pharmacol.* 55 (1975) 221–227.
- [19] W. Marki, M. Brown, J.E. Rivier, Bombesin analogs: effects on thermoregulation and glucose metabolism, *Peptides* 2 (1981) 169–177.
- [20] J.A. Carver, The conformation of bombesin in solution as determined by two-dimensional h-1-nmr techniques, *Eur. J. Biochem.* 168 (1987) 193–199.
- [21] J.A. Carver, A two-dimensional H-1-NMR study of the solution conformation of gastrin releasing peptide, *Biochem. Biophys. Res. Commun.* 150 (1988) 552–560.
- [22] C. DiBello, L. Gozzini, M. Tonellato, M.G. Corradini, G. D'Auria, L. Paolillo, E. Trivellone, 500 MHz NMR characterization of synthetic bombesin and related peptides in DMSO-D6 by two-dimensional techniques, *FEBS Lett.* 237 (1988) 85–90.
- [23] J.A. Carver, J.G. Collins, NMR identification of a partial helical conformation for bombesin in solution, *Eur. J. Biochem.* 187 (1990) 645–650.
- [24] M.D. Dlaz, M. Fioroni, K. Burger, S. Berger, Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: an NMR spectroscopic and molecular dynamics study, *Chem. Eur. J.* 8 (2002) 1663–1669.
- [25] D. Erne, R. Schwyzler, Membrane-structure of bombesin studied by infrared-spectroscopy – prediction of membrane interactions of gastrin-releasing peptide, neuromedin-B, and neuromedin-C, *Biochemistry* 26 (1987) 6316–6319.
- [26] P. Cavatorta, G. Farruggia, L. Masotti, G. Sartor, A.G. Szabo, Conformational flexibility of the hormonal peptide bombesin and its interaction with lipids, *Biochem. Biophys. Res. Commun.* 141 (1986) 99–105.
- [27] P. Sharma, P. Singh, K. Bisetty, F.J. Corcho, J.J. Perez, Conformational profile of bombesin assessed using different computational protocols, *J. Mol. Graph. Model.* 29 (2010) 581–590.
- [28] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 12, University of California, San Francisco, 2012.
- [29] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935.
- [30] D. Hamelberg, J. Mongan, J.A. McCammon, Accelerated molecular dynamics: a

- promising and efficient simulation method for biomolecules, *J. Chem. Phys.* 120 (2004) 11919–11929.
- [31] L.C. Pierce, R. Salomon-Ferrer, C.A.F. de Oliveira, J.A. McCammon, R.C. Walker, Routine access to millisecond time scale events with accelerated molecular dynamics, *J. Chem. Theory Comput.* 8 (2012) 2997–3002.
- [32] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins* 65 (2006) 712–725.
- [33] R. Horst, G. Wider, J. Fiaux, E.B. Bertelsen, A.L. Horwich, K. Wüthrich, Proton–proton Overhauser NMR spectroscopy with polypeptide chains in large structures, *Proc. Natl. Acad. Sci. U.S.A.* 103 (2006) 15445–15450.
- [34] J.J. Kuszewski, R.A. Thottungal, C.D. Schwieters, G.M. Clore, Automated error-tolerant macromolecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments: improved robustness and performance of the PASD algorithm, *J. Biomol. NMR* 41 (2008) 221–239.
- [35] I. Daidone, A. Amadei, Essential dynamics: foundation and applications, *WIREs comput. Mol. Sci.* 2 (2012) 762–770.
- [36] J.J. Perez, M.S. Tomas, J. Rubio-Martinez, Assessment of the sampling performance of multiple-copy dynamics versus a unique trajectory, *J. Chem. Inf. Model.* 56 (2016) 1950–1962.
- [37] J.H. Ward, Hierarchical grouping to optimize an objective function, *J. Am. Stat. Assoc.* 58 (1963) 236–244.
- [38] R.W. Newberry, R.T. Raines, Secondary forces in protein folding, *ACS Chem. Biol.* 14 (2019) 1677–1686.
- [39] M. Cristau, C. Devin, C. Oiry, O. Chaloin, M. Amblard, N. Bernad, A. Heitz, J.-A. Fehrentz, J. Martinez, Synthesis and biological evaluation of bombesin constrained analogues, *J. Med. Chem.* 43 (2000) 2356–2361.
- [40] M. Vincenzi, F.A. Mercurio, M. Leone, About TFE: old and new findings, *Curr. Protein Pept. Sci.* 20 (2019) 425–451.
- [41] B.K. Kobilka, G protein coupled receptor structure and activation, *Biochim. Biophys. Acta* 1768 (2007) 794–807.



## 4. Conclusions

Considering the results of this work as a whole, the main conclusion is that the choice of suitable mathematical methods permits to have more educated descriptions of physical phenomena. On the other hand, it is also demonstrated how the explanation of empirical results push on for the exploration of new limits in diverse areas of mathematics. This thesis is expected to be a contribution for chemical engineering and physical sciences upgrading, validating this way the principle thesis that motivated its own birth: mathematical development is and has always been essential for engineering and physical sciences to achieve new groundbreaking stands.

1. In the first paper and, to a lesser extent, in the neural network training procedure carried out in the third paper, it is proved that local optimization algorithms are less suitable to obtain a global minimum for multivariable problems. Present results permit to conclude that *GlobalSearch* and local optimization algorithms applied iteratively at diverse initial points are more suitable for problems with a high number of parameters and heavy routines, such as the one concerning the first paper, because of the lower impact of the computational cost compared to stochastic global optimization Genetic Algorithm (GA) routine. However, the higher accuracy of GA is preferred when much lighter routines are optimized, since less computational cost is compromised as happens in the second, third and fourth paper.

2. The impact of optimization methods in obtaining reliable results has been proved throughout almost all the works included in this thesis. Nevertheless, the impact is perhaps higher in the first paper, where a novel iterative algorithm consisting of two-level parameter optimization, enables to obtain fourteen parameters, namely six kinetic parameters and eight Binary Interaction Parameters (BIPs) for Peng-Robinson EoS novel in literature, achieving a significant contribution for chemical engineering upgrading.
3. As far as partial differential equations (PDE) solving methods is concerned, orthogonal collocation has proved to be a suitable option for solving nonlinear PDE systems, as reduces the complexity to an ordinary differential equation (ODE) system obtaining accurate results with few collocation points. The resolution by means of orthogonal collocation of the non-linear PDE system presented in the second paper, based on the shrinking-core model and the modelling of multiphase phenomena through discrete solubility values, determines a breakthrough in near-critical extraction modelling.
4. The simpler algorithm of finite differences method and its proper accuracy with the necessary integration step makes it more suitable for solving a single linear PDE. Hence, the original solution of a two coupled linear PDE system by integrating one of them through Laplace transform followed by the application of finite differences method has demonstrated to be suitable for this kind of problems, enabling to obtain accurate results as proved in the fourth paper.



5. The impact for chemical engineering of those two PDE solving algorithms is clear: enabling the resolution of complex physical models that enlighten novel mass transfer parameters and solubility values for lanolin extraction with modified CO<sub>2</sub>. This is a groundbreaking process interesting for the quality of the wool obtained and for the pure lanolin extracted, a natural wax high-valued in chemical industry.
6. The iterative training algorithm of a nonlinear autoregressive neural network designed in the third paper has proved to predict accurately the behavior of the physical phenomena involved, demonstrating its suitability for other complex phenomena hard to model. Moreover, both the network input structure of the time series data under diverse operation conditions and the Weibull statistical model developed for data augmentation might be a real novelty in extraction modelling.
7. The potential benefits of the nonlinear neural network model in the chemical engineering field have been demonstrated, since the model enables the prediction of optimal operating conditions for a desired purpose, such as maximum extraction under certain operating restrictions, with an excellent performance once neural network is already trained.
8. The last paper proves that 4  $\mu$ s molecular dynamics simulation in explicit solvent appears to be a long enough trajectory to sample the conformational space of a tetradecapeptide like bombesin in comparison to previous works. The use of Principal Component Analysis permits to understand the essential fluctuations of the peptide, providing a qualitative picture of the most important valleys and barriers of the Free Energy Landscape of the peptide.

9. The study of the conformational profile of the bombesin peptide reveals that the lowest energy conformation is a helix at the C-terminus. This conformation could be relevant for its recognition to its receptors since it is known that the segment (6-14) bombesin exhibits also affinity. Moreover, the peptide also adopts a hairpin-turn at the N-terminus that can be related to increase its affinity through its interaction with the extracellular loops of its receptor.

## 5. References

- [1] D. M. Burton, *The History of Mathematics. An Introduction*. 7th Edition, New York: McGraw-Hill, 2011.
- [2] M. i. M. al-Khwārizmī, *The Algebra*, F. Rosen, Ed., London: Oriental Translation Fund, 1831.
- [3] F. Spagnolo and B. Di Paola, "European and Chinese Cognitive Styles and Their Impact on Teaching Mathematics," *Studies in Computational Intelligence*, vol. 227, 2010.
- [4] J. N. Crossley and A. S. Henry, "Thus Spake al-Khwsrizmi: A Translation of the Text of Cambridge University Library Ms. Ii.vi.5," *Historia Mathematica*, vol. 17, pp. 103-131, 1990.
- [5] G. Ifrah, *The Universal History of Numbers from Prehistory to the Invention of the Computer*, New York: John Wiley & Sons, Inc., 200.
- [6] Euclid, *Euclid's Elements of Geometry*, R. Fitzpatrick, Ed., 2007.
- [7] R. O. J. Wells, "Geometry in the Age of Enlightenment," *eprint arXiv, Cornell University*, no. 1507.00060, 2015.
- [8] L. Indorato and P. Nastasi, "The 1740 Resolution of the Fermat-Descartes Controversy," *Historia Mathematica*, vol. 16, pp. 137-148, 1989.
- [9] J. I. Carrero-Mantilla, "The numerical problems within analytical methods of solution for cubic equations of state," *Ingeniería y Competitividad*, vol. 14, no. 1, pp. 75 - 89, 2012.
- [10] U. K. Deiters, "Calculation of Densities from Cubic Equations of State," *AICHE Journal*, vol. 48, no. 4, pp. 882-886, 2002.

- [11] A. Valverde, L. Osmieri and F. Recasens, "Binary interaction parameters from reacting mixture data. Supercritical biodiesel process with CO<sub>2</sub> as cosolvent," *The Journal of Supercritical Fluids*, vol. 143, p. 107–119, 2019.
- [12] L. Osmieri, R. A. Moghadam and F. Recasens, "Continuous biodiesel production in supercritical two-step process: phase equilibrium and process design," *The Journal of Supercritical Fluids*, vol. 124, pp. 57-71, 2017.
- [13] S. Saka and Y. Isayama, "A new process for catalyst-free production of biodiesel using supercritical methyl acetate," *Fuel*, vol. 88, pp. 1307-1313, 2009.
- [14] P. Campanelli, M. Banchemo and L. Manna, "Synthesis of biodiesel from edible, non-edible and waste cooking oils via supercritical methyl acetate transesterification," *Fuel*, vol. 89, pp. 3675-3678, 2010.
- [15] J. Maçaira, A. Santana, F. Recasens and M. A. Larrayoz, "Biodiesel production using supercritical methanol/carbon dioxide mixtures in a continuous reactor," *Fuel*, vol. 90, pp. 2280-2288, 2011.
- [16] G. Brunner, Gas extraction, An Introduction to Fundamentals of Supercritical Fluids and the Application to Separation Processes, Darmstadt-New York,,: Springer Verlag, 1994.
- [17] J.-N. Jaubert and L. Coniglio, "The group contribution concept: a useful tool to correlate binary systems and to predict the phase behavior of multicomponent systems involving supercritical CO<sub>2</sub> and fatty acids," *Industrial & Engineering Chemistry Research*, vol. 38, p. 5011–5018, 1999.
- [18] R. Reid, Supercritical-fluid extraction. A perspective, University of Wisconsin: Hougen Lecture Series, 1981.

- [19] O. Pfohl, S. Petkhov and G. Brunner, *A Powerful Tool to Correlate Phase Equilibria*, München: Herbert Utz Verlag, 2000.
- [20] X. Xu, J.-N. Jaubert, R. Privat and P. Arpentier, "Prediction of thermodynamic properties of alkane-containing mixtures with the E-PPR78 model," *Industrial & Engineering Chemistry Research*, vol. 56, p. 8143–8157, 2017.
- [21] H. Gros, E. Bottini and E. Brignole, "Group contribution equation of state for associating mixtures," *Fluid Phase Equilibria*, vol. 116, p. 537–544, 1996.
- [22] A. Bertucco, P. Canu, L. Devetta and A. Zwahlen, "Catalytic hydrogenation in supercritical critical CO<sub>2</sub>: kinetic measurements in a gradientless internal-recycle reactor," *Industrial & Engineering Chemistry Research*, vol. 36, pp. 2626-2633, 1997.
- [23] E. Santacesaria, M. Di Serio, R. Tesser and M. Cozzolino, "Prediction of VLE for the Kinetic Study of High-pressure Processes Based on Synthesis Gas," in *DGMK/SCI Conference "Synthesis Gas Chemistry" DGMK Tagungsbericht 2006-4*, DGMK, Dresden, 2006.
- [24] M. Nauenberg, "The Reception of Newton's Principia," *eprint arXiv, Cornell University*, no. 1503.06861, 2015.
- [25] B. Pourciau, "Newton and the Notion of Limit," *Historia Mathematica*, vol. 28, p. 18–30, 2001.
- [26] I. Newton, *Philosophiæ Naturalis Principia Mathematica*, London, 1686.
- [27] C. A. Lehner and H. Wendt, "Mechanics in the Querelle des Anciens et des Modernes," *Isis*, vol. 108, no. 1, pp. 26-39, 2017.
- [28] F. Cajori, "The History of Notations of the Calculus," *Annals of Mathematics, Second Series*, vol. 25, no. 1, pp. 1-46, 1923.

- [29] R. Thom, "The Role of Mathematics in Present-day Science," in *Logic, Methodology and Philosophy of Science VI, Proceedings of the Sixth International Congress of Logic, Methodology and Philosophy of Science*, L. J. C. Jerzy, Ed., Hannover, PWN-Polish Scientific Publishers-Warszawa and North-Holland Publishing Company, 1979, pp. 3-13.
- [30] G. L. Alexanderson, "About the cover: Euler and Königsberg's Bridges: A historical view," *Bulletin (New Series) of the American Mathematical Society*, vol. 43, no. 4, p. 567–573, 2006.
- [31] E. Hairer, G. Wanner and S. P. Norsett, *Solving Ordinary Differential Equations I, Nonstiff Problems*, Berlin: Springer-Verlag, 1993.
- [32] T. Kimura, "On Dormand-Prince Method," 24 September 2009. [Online]. Available:  
[http://depa.fquim.unam.mx/amyd/archivero/DormandPrince\\_19856.pdf](http://depa.fquim.unam.mx/amyd/archivero/DormandPrince_19856.pdf).  
[Accessed 12 February 2020].
- [33] J. E. Márquez Díaz, "Fifth postulate of Euclid and the non-Euclidean geometries. Implications with the spacetime," *International Journal of Scientific and Engineering Research*, vol. 9, no. 3, pp. 530-542, 2018.
- [34] W. H. Bussey, "Non-Euclidean Geometry," *The Mathematics Teacher*, vol. 15, no. 8, pp. 445-459, 1922.
- [35] J. Gray, "Gauss and Non-Euclidean Geometry," in *Non-Euclidean Geometries. Mathematics and Its Application.*, vol. 581, Boston, MA: Springer, 2006, pp. 61-80.
- [36] H. Kragh, "Geometry and Astronomy: Pre-Einstein Speculations of Non-Euclidean Space," *eprints arXiv, Cornell University*, no. 1205.4909, 2012.

- [37] A. M. Naveira, "The Riemann Curvature Through History," *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, vol. 99, no. 2, p. 195–210, 2005.
- [38] C. Lämmerzahl, "Relativity and technology," *Annals of Physics*, vol. 15, no. 1 – 2, 5 – 18, 2006.
- [39] D. Dieks, "The Flexibility of Mathematics," in *The Role of Mathematics in Physical Sciences*, Dordrecht, Springer, 2005, pp. 115-129.
- [40] P. Rebentrost, M. Schuld, L. Wossnig, F. Petruccione and S. Lloyd, "Quantum gradient descent and Newton's method for constrained polynomial optimization," *preprints arXiv, Cornell University*, no. 1612.01789, 2018.
- [41] G. Becigneul and O.-E. Ganea, "Riemannian Adaptive Optimization Methods," in *The International Conference on Learning Representations*, New Orleans, 2019.
- [42] A. R. Conn, N. I. M. Gould and P. L. Toint, *Trust Region Methods*, Philadelphia: Mathematical Programming Society - Society for Industrial and Applied Mathematics, 2000.
- [43] Y.-x. Yuan, "A review of trust region algorithms for optimization," in *International Congress on Industrial and Applied Mathematics*, 2000.
- [44] S. A. Santos, "Trust-Region-based Methods for Nonlinear Programming: Recent Advances and Perspectives," *Pesquisa Operacional*, vol. 34, no. 3, 2014.
- [45] P.-A. Absil, C. G. Baker and K. A. Gallivan, "Trust-Region Methods on Riemannian Manifolds," *Foundations of Computational Mathematics*, vol. 7, no. 3, pp. 303-330, 2007.

- [46] J. Mentel, V. Tiller, E. Moller and D. Haberland, "Estimation of parameters in systems of ordinary differential-equations to the determination of kinetic parameters," *Chemische Technik*, vol. 44, no. 9, pp. 300-303, 1992.
- [47] A. M. Almagrbi, T. Hatami, S. B. Glisic and A. M. Orlović, "Determination of kinetic parameters for complex transesterification reaction by standard optimisation methods," *Hemijska Industrija*, vol. 68, p. 149–159, 2014.
- [48] T. Weise, "Global Optimization Algorithms - Theory and Application," 2009.
- [49] MATLAB®, "Global Optimization Toolbox User's Guide," The Math Works Inc., Natick, 2016.
- [50] C. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, London: John Murray, 1859.
- [51] X.-S. Yang, "Optimization and Metaheuristic Algorithms in Engineering," in *Metaheuristics in Water, Geotechnical and Transport Engineering*, Oxford, Elsevier, 2013, pp. 1-23.
- [52] A. Meyer-Baese and V. Schmid, "Genetic Algorithms," in *Pattern Recognition and Signal Analysis in Medical Imaging*, Oxford, Academic Press - Elsevier, 2014, p. 135–149.
- [53] J. S. Arora, "Genetic Algorithms for Optimum Design," in *Introduction to Optimum Design*, Oxford, Academic Press - Elsevier, 2012, p. 643–655.
- [54] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 7, p. 115–133, 1943.



- [55] G. Piccinini, "The First Computational Theory of Mind and Brain: A Close Look at McCulloch and Pitts's "Logical Calculus of Ideas Immanent in Nervous Activity'," *Synthese*, vol. 141, p. 175–215, 2004.
- [56] L. A. Gladwin, "Alan Turing, Enigma, and the Breaking of German Machine Ciphers in World War II," *Pologue*, vol. 29, no. 3, pp. 203-217, 1997.
- [57] R. G. M. Morris, "D.O. Hebb: The Organization of Behavior, Wiley: New York; 1949," *Brain Research Bulletin*, vol. 50, no. 5/6, p. 437, 1999.
- [58] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Technical Report IDSIA*, no. arXiv:1404.7828 v4, 2014.
- [59] M. Girardi-Schappo, M. H. R. Tragtenberg and O. Kinouchi, "A Brief History of Excitable Map-Based Neurons and Neural Networks," *Journal of Neuroscience Methods*, vol. 220, no. 2, pp. 116-130, 2013.
- [60] I. N. Da Silva, D. Hernane Spatti, R. Andrade Flauzino, L. H. B. Liboni and S. F. dos Reis Alves, "Liboni, L. H. B., & dos Reis Alves, S. F.. Artificial Neural Network Architectures and Training Processes," in *Artificial Neural Networks*, Basel, Springer, 2016, p. 21–28.
- [61] L. S. H. Ngia and J. Sjöberg, "Efficient Training of Neural Nets for Nonlinear Adaptive Filtering Using a Recursive Levenberg–Marquardt Algorithm," *IEEE Transaction on Signal Processing*, vol. 48, no. 7, pp. 1915-1927, 2000.
- [62] A. Valverde, J. Alvarez-Florez and F. Recasens, "Hybrid Nonlinear Autoregressive Neural Network – Weibull Statistical Model Applied to the Supercritical Extraction of Lanolin from Raw Wool," *Submitted to SN Applied Sciences (03/04/2020). Under Review*, 2020.

- [63] M. Cygnarowicz-Provost, J. W. King, W. M. Marmer and P. Magidman, "Extraction of Woolgrease with Supercritical Carbon Dioxide," *Journal of the American Oil Chemists' Society*, vol. 71, p. 222–225, 1994.
- [64] B. S. Kuo, J. C. Kim, J. H. Jeon and H. K. Bae, "Desorption of Wool Grease from Greasy Wool with Supercritical Carbon Dioxide," *Hwahak Konghak*, vol. 30, pp. 491-498, 1992.
- [65] F. W. Jones, D. R. Bateup, S. R. Dixon and S. R. Gray, "Solubility of Wool Wax in Supercritical Carbon Dioxide," *The Journal of Supercritical Fluids*, vol. 10, pp. 100-111, 1995.
- [66] V. Eychenne, S. Sáiz, F. Trabelsi and F. Recasens, "Near-critical solvent extraction of wool with modified carbon dioxide - experimental results," *The Journal of Supercritical Fluids*, vol. 21, pp. 23-31, 2001.
- [67] M. Fullana, F. Trabelsi and F. Recasens, "Use of neural net computing for statistical and kinetic modelling and simulation of supercritical fluid extractors," *Chemical Engineering Science*, vol. 54, pp. 5845-5862, 1999.
- [68] A. B. Bulsari, *Neural networks for chemical engineers*, Amsterdam: Elsevier, 1995.
- [69] A. B. Bulsari and S. Palosaari, "Application of Neural Networks for System Identification of an Adsorption Column," *Neural Computing and Applications*, vol. 1, pp. 160-165, 1993.
- [70] C.-D. Lai, "Weibull Distributions and Their Applications," in *Springer Handbook of Engineering Statistics*, New York, Springer, 2006, pp. Party A-2, 1-16.
- [71] T. D. Chaudhuri and I. Ghosh, "Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate

- Framework," *Journal of Insurance and Financial Management*, vol. 1, no. 5, pp. 92-123, 2016.
- [72] T. Lin, B. Horne, P. Tino and C. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *IEEE Transactions on Neural Networks*, vol. 7, no. 6, p. 1329–1338, 1996.
- [73] K. Lang, A. Waibel and G. E. Hinton, "A time-delay neural network architecture for isolated word recognition," *Neural Networks*, vol. 3, p. 23–43, 1990.
- [74] "Nonlinear Autoregressive Neural Network Models for Prediction of Transformer Oil-Dissolved Gas Concentrations," *Energies*, vol. 11, no. 1691, 2018.
- [75] K. R. Davidson and A. P. Donsig, "Fourier Series and Physics," in *Real Analysis and Applications*, Berlin, Springer, 2009, p. 328–359.
- [76] I. Grattan-Guinness, "Joseph Fourier and the Revolution in Mathematical Physics," *Journal of the Institute of Mathematics and its Applications*, vol. 5, pp. 230-253, 1969.
- [77] C. Roychoudhuri, "Bi-centenary of successes of Fourier theorem: its power and limitations in optical system designs," *Current Developments in Lens Design and Optical Engineering VIII*, vol. 6667, no. 66670D, 2007.
- [78] B. E. Rapp, "Transforms," in *Microfluidics: Modelling, Mechanics and Mathematics*, Karlsruhe, Elsevier, 2017, p. 81–92.
- [79] S. H. Lin and H. Eyring, "Solving the time-dependent Schrödinger equation via Laplace Transform," *Proceedings of the National Academy of Sciences*, vol. 68, no. 1, pp. 76-81, 1971.

- [80] N. Riahi, "Solving the time-dependent Schrödinger equation via Laplace transform," *Quantum Studies: Mathematics and Foundations*, vol. 4, p. 103–126, 2017.
- [81] E. C. Molina, "The theory of probability: some comments on Laplace's *théorie analytique*," *Bulletin of the American Mathematical Society*, vol. 36, no. 6, pp. 369-392, 1930.
- [82] A. Valverde, J. Alvarez-Florez and F. Recasens, "Mathematical modelling of supercritical fluid extraction of liquid lanolin from raw wool. Solubility and mass transfer rate parameters," *Submitted to Chemical Engineering Research and Design (11/12/2019). Under Review*, 2020.
- [83] J. Puiggené, M. A. Larrayoz and F. Recasens, "Free liquid-to-supercritical fluid mass transfer in packed beds," *Chemical Engineering Science*, vol. 52, no. 2, pp. 195-212, 1997.
- [84] B. Jain and A. Sheng, "An Explor An Exploration of the Appr ation of the Approximation of Deriv ximation of Derivative Functions via unctions via Finite Differences," *Rose-Hulman Undergraduate Mathematics Journal*, vol. 8, no. 2, 2007.
- [85] D. Hilbert, *Grundlagen der Geometrie*. First edition, Göttingen: Tuebner, 1899.
- [86] M. F. Rañada, "David Hilbert, Hermann Minkowski, y la Axiomatización de la Física," *La Gaceta de la Real Sociedad Matematica Española*, vol. 6.3, p. 641–664, 2003.
- [87] G. Moroşanu, *Functional Analysis for the Applied Sciences*, Bucharest: Springer, 2019.

- [88] S. S. Holland, *Applied Analysis by the Hilbert Space Method: An Introduction with Applications to the Wave, Heat, and Schrödinger Equations*, Mineola: Dover Publications, Inc, 2012.
- [89] K. Landsman, "Quantum theory and functional analysis," in *To appear in the Oxford Handbook of the History of Interpretations and Foundations of Quantum*, Oxford, Oxford University Press, 2021.
- [90] M. Rédei, "John Von Neumann on Mathematical and Axiomatic Physics," in *The Role of Mathematics in Physical Sciences. Interdisciplinary and Philosophical Aspects*, Dordrecht, Springer, 2005, pp. 43-54.
- [91] K. S. Ranade, "Functional analysis and quantum mechanics: an introduction," *Fortschritte der Physik*, vol. 63, no. 9-10, p. 644–658, 2015.
- [92] K. Ciesielski, "On Stefan Banach and Some of his Results," *Banach Journal of Mathematical Analysis*, vol. 1, no. 1, pp. 1-10, 2007.
- [93] L. A. Steen, "Highlights in the history of spectral theory," *American Mathematical Monthly*, vol. 80, p. 359–381, 1973.
- [94] E. R. Tou, "Math Origins: Eigenvectors and Eigenvalues," *Convergence*, 2018.
- [95] D. Hinton, B. Simon, W. N. Everitt, J. Weidmann, Y. Last, D. Gilbert, C. Bennewitz, V. A. Galaktionov, C.-N. Chen, R. Róo and M. M. Malamud, *Sturm-Liouville Theory*, W. O. Amrein, A. M. Hinz and D. P. Pearson, Eds., Berlin: Springer, 2005.
- [96] S. Bravo Yuste, *Métodos matemáticos avanzados para científicos e ingenieros*. Colección manuales uex - 48, Cáceres: Universidad de Extremadura. Servicio de Publicaciones, 2006.

- [97] E. Schrödinger, "Quantisierung als Eigenwertproblem," *Annalen der Physik*, vol. 79, p. 361–376, 1926.
- [98] D. A. B. Miller, *Quantum Mechanics for Scientists and Engineers*, Cambridge: Cambridge University Press, 2008.
- [99] D. A. Bonilla, A. Martín-Ruiz and L. F. Urrutia, "Exact solution of the Schrödinger equation for a hydrogen atom at the interface between the vacuum and a topologically insulating surface," *The European Physical Journal D*, vol. 73, no. 114, 2019.
- [100] J. V. Villadsen and W. E. Stewart, "Solution of boundary-value problems by orthogonal collocation," *Chemical Engineering Science*, vol. 22, p. 1483–1501, 1967.
- [101] A. S. Olagunju, L. J. Folake and M. T. Raji, "Comparative study of the effect of different collocation points on Legendre-collocation methods of solving second-order boundary value problems," *IOSR Journal of Mathematics*, vol. 7, p. 35–41, 2013.
- [102] B. A. Finlayson, "Packed bed reactor analysis by orthogonal collocation," *Chemical Engineering Science*, vol. 1081–1091, p. 26, 1971.
- [103] M. Morbidelli, A. Servida, G. Storti, R. Paludetto and S. Carrà, "Application of the orthogonal collocation method to some chemical engineering problems," *Gazzetta chimica Italiana*, vol. 19, pp. 47-60, 1983.
- [104] A. Valverde and F. Recasens, "Extraction of solid lanolin from raw wool with near-critical ethanol-modified CO<sub>2</sub>—A mass transfer model," *The Journal of Supercritical Fluids*, vol. 145, pp. 151-161, 2019.

- [105] O. Levenspiel, "Chapter 25. Fluid-Particle Reactions," in *Chemical Reaction Engineering. Third Edition*, New York, John Wiley & Sons, 1999, pp. 566-588.
- [106] S. Ashley, "Core Concept: Ergodic theory plays a key role in multiple fields," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 112, no. 7, p. 1914, 2015.
- [107] J. v. Plato, "Boltzmann's ergodic hypothesis," *Archive for History of Exact Sciences*, vol. 42, p. 71–89, 1991.
- [108] P. Ehrenfets and T. Ehrenfest, "Begriffliche Grundlagen der statistischen Auffassung in der Mechanik," in *Enzyklopädie der Mathematischen Wissenschaften vol. 4*, Leipzig, Teubner, 1911.
- [109] J. C. Maxwell, "Illustrations of the Dynamical Theory of Gases. The Kinetic Theory of Gases," *Philosophical Magazine*, Vols. 19-20, p. 19–32; 21–37, 1860.
- [110] G. Gallavotti, "Ergodicity: a historical perspective. Equilibrium and Nonequilibrium," *The European Physical Journal H*, vol. 41, pp. 181-259, 2016.
- [111] L. Corry, "David Hilbert and the Axiomatization of Physics (1894-1905)," *Archive for History of Exact Sciences*, vol. 51, no. 2, pp. 83-198, 1997.
- [112] U. Cárcamo, "El origen fenomenológico de la teoría ergódica," *Revista Universidad EAFIT*, vol. 32, no. 103, pp. 15-27., 2012.
- [113] L.-Q. Yang , X.-L. Ji and S.-Q. Liu, "The free energy landscape of protein folding and dynamics: a global view," *Journal of Biomolecular Structure and Dynamics*, vol. 31, no. 9, pp. 982-992, 2013.
- [114] J. M. Yon, "Protein folding: a perspective for biology, medicine and biotechnology," *Brazilian Journal of Medical and Biological Research*, vol. 34, pp. 419-435, 2001.

- [115] R. Zwanzig, A. Szabo and B. Bagchi, "Levinthal's paradox," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 89, pp. 20-22, 1992.
- [116] C. Oostenbrink, M. M. H. v. Lipzig and W. F. v. Gunsteren, "Applications of Molecular Dynamics Simulations in Drug Design," in *Comprehensive Medicinal Chemistry II. Volume 4*, Zürich, Elsevier, 2007, pp. 651-668.
- [117] J. A. McCammon, B. R. Gelin and M. Karplus, "Dynamics of folded proteins," *Nature*, vol. 267, p. 585-590, 1977.
- [118] S. A. Hollingsworth and R. O. Dror, "Molecular Dynamics Simulation for All," *Neuron*, vol. 99, pp. 1129-1143, 2018.
- [119] A. T. P. Carvalho, A. Barrozo, D. Doron, A. V. Kilshtain, D. T. Major and S. C. L. Kamerlin, "Challenges in computational studies of enzyme structure, function and dynamics," *Journal of Molecular Graphics and Modelling*, vol. 54, p. 62-79, 2014.
- [120] G. Stock, A. Jain, L. Riccardi and P. H. Nguyen, "Exploring the Energy Landscape of Small Peptides and Proteins by Molecular Dynamics Simulations," in *Protein and Peptide Folding, Misfolding, and Non-Folding. First Edition*, New York, John Wiley & Sons, Inc., 2012, pp. 57-77.
- [121] I. Daidone and A. Amadei, "Essential dynamics: foundation and applications," *WIREs Computational Molecular Science*, vol. 00, pp. 1-9, 2012.
- [122] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A*, vol. 374, 2016.



- [123] A. Valverde, P. Gomez-Gutierrez and J. J. Perez, "Assessment of the Conformational Profile of Bombesin by Computational Methods," *Journal of Molecular Graphics and Modelling*, vol. 98, no. 107590, 2020.
- [124] H. C. Weber, "Regulation and signaling of human bombesin receptors and their biological effects," *Current Opinion in Endocrinology, Diabetes and Obesity*, vol. 16, pp. 66-71, 2009.
- [125] R. T. Jensen, J. F. Battey, E. R. Spindel and R. V. Beny, "International Union of Pharmacology. LXVIII. Mammalian Bombesin Receptors: Nomenclature, Distribution, Pharmacology, Signaling, and Functions in Normal and Disease States," *Pharmacological Reviews*, vol. 60, pp. 1-42, 2008.
- [126] V. Erspamer and P. Melchiorri, "Active polypeptides - from amphibian skin to gastrointestinal-tract and brain of mammals," *Trends in Pharmacological Sciences*, vol. 1, pp. 391-395, 1980.
- [127] R. T. Jensen and T. W. Moody, "Bombesin peptides (cancer)," in *Hand-book of biologically active peptides*, Amsterdam, Elsevier, 2013, p. 506–511.
- [128] J. A. Carver and J. G. Collins, "NMR identification of a partial helical conformation for bombesin in solution," *European Journal of Biochemistry*, vol. 187, pp. 645-650, 1990.
- [129] M. D. Díaz, M. Fioroni, K. Burger and S. Berger, "Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: An NMR spectroscopic and molecular dynamics study," *Chemistry – A European Journal*, vol. 8, pp. 1663-1669, 2002.

- [130] P. Sharma, P. Singh, K. Bisetty, F. J. Corcho and J. J. Pérez, "Conformational profile of bombesin assessed using different computational protocols," *Journal of Molecular Graphics and Modelling*, vol. 29, p. 581–590, 2010.
- [131] M. S. Mahoney, "The History of Computing in the History of Technology," *Annals of the History of Computing*, vol. 10, pp. 113-125, 1998.
- [132] M. P. Domínguez, "ENIAC, matemáticas y computación científica," *Gaceta de la Real Sociedad Matemática Española*, vol. 2, no. 3, pp. 495-518, 1999.
- [133] J. Molero Prieto, "Del ENIAC, hasta los andares," *ReVisión*, vol. 7, no. 1, pp. 35-51, 2014.
- [134] J. Backus, "The History of FORTRAN I, II and III," in *History of Programming Languages*, New York, Academic Press, Inc., 1981, pp. 25-74.
- [135] D. Ritchie, "The development of the C language," in *History of Programming Languages II (HOPL-II)*, Cambridge, Massachusetts, 1993.
- [136] E. C. D.-G. f. Research, "Research.eu," 2009. [Online]. Available: [https://ec.europa.eu/research/audio/women-in-science/pdf/wis\\_en.pdf](https://ec.europa.eu/research/audio/women-in-science/pdf/wis_en.pdf). [Accessed 09 03 2020].
- [137] L. A. Caicedo, G. Rodríguez and A. Durán, "Las Matemáticas y La Ingeniería Química, Una Relación Sinérgica," *Ingeniería e Investigación*, vol. 45, pp. 47-52, 2000..
- [138] D. Bogoya, "Naturaleza de la Simulación," *Ingeniería e Investigación*, vol. 29, p. 45, 1996.
- [139] J. W. Mellor, *Higher Mathematics for Students of Chemistry and Physics with Special Reference to Practical Work*, London: Longmans, Green & Co., 1902.

- [140] R. S. Carslaw , *Fourier Series and Integrals and Mathematical Theory of the Conduction Heat*, London: Macmillan, 1906.
- [141] R. B. Bird, W. E. Stewart and E. N. Lightfoot, *Transport Phenomena*, New York: Wiley, 1960.
- [142] A. G. Fredrickson, "Stochastic Triangular Reactions," *Chemical Engineering Science*, vol. 21, pp. 687-691, 1966.
- [143] L. T. Fan, J. R. Too and R. Nassar, "Stochastic simulation of residence time distribution curves," *Chemical Engineering Science*, vol. 40, no. 9, pp. 1743-1749, 1985.
- [144] J. M. Smith, *Chemical Engineering Kinetics*, Third edition, New York: McGraw-Hill, 1981.
- [145] R. J. Farrauto and C. H. Bartholomew, *Fundamentals of Industrial Catalytic Processes*, London: Chapman and Hall, 1997.
- [146] K. Abaroudi, F. Trabelsi and F. Recasens, "Screening of cosolvents for a supercritical fluid. A fully predictive approach," *AIChE Journal*, vol. 48, p. 551–560, 2002.
- [147] J. M. Bayona, P. Erra, Z. Moldovan, C. Domínguez, E. Jover, F. Recasens and M. A. Larrayoz, "Method for obtaining lipid fractions from wool or lanolin using pressurized carbon dioxide". Oficina Española de Patentes y Marcas, Madrid Patent WO 2002100990A1, 16 07 2005.
- [148] C.-S. Tan, S.-K. Liang and D.-C. Liou, "Fluid-solid mass transfer in a supercritical fluid extractor," *Chemical Engineering Journal*, vol. 38, pp. 17-23, 1980.

- [149] M. B. King, T. R. Bott, M. J. Barr and R. S. Mahmud, "Equilibrium and rate data for the extraction of Lipids using compressed carbon dioxide," *Separation Science and Technology*, vol. 22, no. 2, p. 1103, 1987.
- [150] I. M. Abdulagatov, N. G. Polikhronidi, A. Abdurashidova, S. B. Kiselev and J. F. Ely, "Thermodynamic Properties of Methanol in the Critical and Supercritical Regions," *International Journal of Thermophysics*, vol. 26, no. 5, 2005.
- [151] A. F. M. Barton, *CRC Handbook of Solubility Parameters and Other Cohesion Parameters*, Boca Ratón: CRC Press LLC, 1991.
- [152] D. W. Green and R. H. Perry, *Perry's Chemical Engineers' Handbook*, New York: McGraw-Hill, 2008.
- [153] Z. Tang, Z. Dua, E. Mina, L. Gaob, T. Jiang and B. Hanb, "Phase equilibria of methanol–triolein system at elevated temperature and pressure," *Fluid Phase Equilibria*, vol. 8, p. 239, 2006.
- [154] W. Weber, S. Petkov and G. Brunner, "Vapour–liquid-equilibria and calculations using the Redlich–Kwong–Aspen-equation of state for tristearin, tripalmitin, and triolein in CO<sub>2</sub> and propane," *Fluid Phase Equilibria*, vol. 695, pp. 158-160, 1999.
- [155] L. Vázquez, A. M. Hurtado-Benavides, G. Reglero, T. Fornari, E. Ibáñez and F. J. Señoráns, "Deacidification of olive oil by countercurrent supercritical carbon dioxide extraction: Experimental and thermodynamic modeling," *Journal of Food Engineering*, vol. 463, p. 90, 2009.
- [156] S. Glisic and D. Skala, "The prediction of critical parameters for triolein, diolein, monoolein and methyl esters," in *9th International Symposium on Supercritical Fluids 2009, New Trends in Supercritical Fluids: Energy, Materials, Processing*, Archon, 2009.

- [157] C. Chlev and E. Simeonov, "Simulation of biodiesel production by transesterification of vegetable oils," *Journal of Chemical Technology and Metallurgy*, vol. 49, no. 5, pp. 479-486, 2014.
- [158] A. Gómez, "Solubility of triolein in dense gases," Maribor, 2011.
- [159] A. C. Petcu, V. Plesu and C. Berbente, "Estimation methods for thermophysical properties of camelina sativa crude oil," *Scientific Bulletin - University Politehnica of Bucharest*, vol. 78, no. 1, 2016.
- [160] F. A. Philip, "Simulation study of distillation, stripping, and flash technology for an energy efficient methanol recovery unit in biodiesel production processes," Regina, 2013.
- [161] R. C. Reid, J. M. Prausnitz and B. E. Poling, *The properties of gases & liquids*, New York: McGraw-Hill, Inc., 1987.
- [162] T. Fornari, D. Tenllado, C. Torres and G. Reglero, "Supercritical Phase Equilibria Modeling of Glyceride Mixtures," *Journal of Thermodynamics*, vol. 2011, p. 9, 2011.
- [163] C. Baroi, "Biodiesel production using supported 12-tungstophosphoric acid as solid acid catalysts," Saskatoon, Canada, 2014.
- [164] A. E. Costa, A. Santana, M. B. Quadri, R. A. F. Machado, F. Recasens and M. A. Larrayoz, "Glycerol desorption from ion exchange and adsorbent resin using supercritical fluid technology: An optimization study," *The Journal of Supercritical Fluids*, vol. 58, pp. 226-232, 2011.
- [165] V. I. Anikeev, "Thermodynamics of phase and chemical equilibrium in the processes of biodiesel fuel synthesis in sub- and supercritical methanol," Novosibirsk, 2012.

- [166] J. W. Picou, "Glycerin reformation in high temperature and pressure water," Doctoral Dissertations, 2012.
- [167] R. P. Singh and A. G. Medina, Food Properties and Computer-Aided Engineering of Food Processing Systems, Dordrecht: Kluwer Academic Publishers, 1989.
- [168] D.-F. Ruan, Z.-H. Chen, K.-F. Wang, Y. Chen and F. Yang, "Physical Property Prediction for Waste Cooking Oil Biodiesel," *The Open Fuels & Energy Science Journal*, vol. 7, pp. 62-68, 2014.
- [169] A. M. A. Karim, A. K. Mutlag and M. S. Hameed, "Vapor-liquid equilibrium prediction by pe and ann for the extraction of unsaturated fatty acid esters by supercritical CO<sub>2</sub>," *Asian Research Publishing Network. Journal of Engineering and Applied Sciences*, vol. 6, no. 9, 2011.
- [170] H. An, W. M. Yang, A. Maghbouli, S. K. Chou and K. J. Chua, "Detailed physical properties prediction of pure methyl esters for biodiesel combustion modeling," *Applied Energy*, vol. 102, p. 647–656, 2013.
- [171] V. Ferioiu, S. Sima and D. Geana, "High pressure phase equilibrium in carbon dioxide+ethanol system," *Scientific Bulletin - University Politehnica of Bucharest*, vol. 75, no. 1, 2013.
- [172] N. Budisa and D. Schulze-Makuch, "Supercritical Carbon Dioxide and Its Potential as a Life-Sustaining Solvent in a Planetary Environment," *Life*, vol. 4, pp. 331-340, 2014.
- [173] E. Goos, U. Riedel, L. Zhao and L. Blum, "Phase diagram of CO<sub>2</sub> and CO<sub>2</sub> -N<sub>2</sub> gas mixtures and their application in compression processes," *Energy Procedia*, vol. 4, pp. 3778-3785, 2011.

- [174] Z. Tang, Z. Du, E. Min, L. Gao, T. Jiang and B. Han, "Phase equilibria of methanol–triolein system at elevated temperature and pressure," *Fluid Phase Equilibria*, vol. 239, p. 8–11, 2006.
- [175] S. Arvelos, L. L. Rade, E. O. Watanabe, C. E. Hori and L. L. Romanielo, "Evaluation of different contribution methods over the performance of Peng–Robinson and CPA equation of state in the correlation of VLE of triglycerides, fatty esters and glycerol + CO<sub>2</sub> and alcohols," *Fluid Phase Equilibria*, vol. 362, p. 136–146, 2014.
- [176] Z. K. Lopez-Castillo, S. N. V. K. Aki, M. A. Stadtherr and J. F. Brennecke, "Enhanced Solubility of Hydrogen in CO<sub>2</sub>-Expanded Liquids," *Industrial & Engineering Chemistry Research*, vol. 47, pp. 570-576, 2008.
- [177] T. J. Hughes, M. E. Kandil, B. F. Graham and E. F. May, "Simulating the capture of CO<sub>2</sub> from natural gas: new data and improved models for methane + carbon dioxide + methanol," *International Journal of Greenhouse Gas Control*, vol. 31, pp. 121-127, 2014.
- [178] G. J. Chin, Z. H. Chee, W. Chen and A. Rajendran, "Solubility of Flurbiprofen in CO<sub>2</sub> and CO<sub>2</sub> + Methanol," *Journal of Chemical & Engineering Data*, vol. 55, p. 1542–1546, 2010.
- [179] S.-S. You, C. S. Lee and K.-P. Yoo, "Modeling of Supercritical-Fluid Phase-Equilibria Using A New Nonrandom Lattice-Fluid Theory," *The Journal of Supercritical Fluids*, vol. 6, pp. 69-84, 1993.
- [180] Y. Medina-Gonzalez, T. Tassaing, S. Camy and J.-S. Condoret, "Phase equilibrium of the CO<sub>2</sub>/glycerol system: Experimental data by in situ FT-IR

spectroscopy and thermodynamic modeling," *Journal of Supercritical Fluids*, vol. 73, pp. 97-107, 2013.

- [181] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1970.
- [182] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in Fortran 77*, Cambridge University Press, 1993.
- [183] V. I. Arnold, G. Boniolo, P. Budinich, L. Crivellari, M. Dorato, D. Dieks, G. Ghirardi, G. Giorello, V. Gómez Pin, M. Rédei, A. Rivadulla, N. Singh, C. Sinigaglia, M. Stöltzner, Z. Šikic and N. Zovko, *The Role of Mathematics in Physical Sciences*, Dordrecht: Springer, 2005.



## Annex I

### On critical properties from literature

**Table 1. Methanol critical properties. Authors without both year and reference have been taken from Abdulagatov et al. [150].**

Author	Tc (K)	Pc (bar)	Author	Tc (K)	Pc (bar)	w
Hannay	505,9	73,82	Eubank	512,78	79,84	
Nedejdine	506,2	70,65	Costello & Bowden	513,15		
Ramsay & Young	513,2	79,59	Ambrose & Townsend	512,64		
de Heen	536,2		Loktev	513	79,74	
Schmidt	513,45 15,2		Yerlett & Wormald	512,6		
Centnerszwer	513,4		Ambrose & Walton	512,64	80,92	
Crismer	513,7		Donham	512,58	80,97	
Young	513,2	79,54	Harrison & Gammon		81,04	
Salwedel	513,2	97,01	Kozlov	512,6	81,04	
Fischer & Reichel	513,8		Simonson et al.	512,6	78,2	
Golik et al.	505,2		Polikhronidi et al.	512,78		
Kay & Donham	512,6	80,97	Kobe & Lynn,	513,15	79,54	
Nozdrev	513,2		Kudchadker et al.	512,58	80,96	
Krichevskii et al.	513,2		Wilhoit & Zwolinski	512,6	80,96	
McCracken et al.	513,2	79,57	IUPAC	512,6	81,04	
Mocharnyuk	513,9		Goodwin	512,56	80,95	
Skaates & Kay	512,68	80,94	Smith & Srivastava	512,6	80,92	
Efremov	513	79,74	Zubarev et al.	512,61	81,03	
Zubarev & Bagdonas	512,7	81,04	Craven & de Reuck	512,6	80,97	
Marshall & Jones	514,7		Bashirov	512,64	80,96	
Kay & Khera	512,32	80	Suleimanov	512,7		
Francesconi et al.	512,6	81	Bazaev et al.	512,78	81	
Brunner et al.	512,43	80,72	Abdulagatov & Ely	512,78		
de Loos et	512,5	80,6	Aspen Plus v9, 2016	512,5	80,84	0,566
Lydersen & Tsochev	512,5	80,6	Barton, 1991 [151]			0,556
Gude & Teja	512,5	80,84	Perry, 2008 [152]	512,7	81,4	0,566

**Table 2. Triolein critical properties.**

Author	Tc (K)	Pc (bar)	w
Tang et al., 2006 [153]	954,1	3,602	1,6862
Weber et al., 1999 [154]	947,1	4,682	1,6862
Vázquez et al, 2004 [155]	1043,3	4,57	
Glisic & Skala, 2009 [156]	977,88	3,34	1,9782
Aspen Plus v9, 2016	1640	4,7	
Glisic & Skala (Aspen) [156]	998	3,34	1,9782
Chilev & Simeonov, 2014 [157]	1366,85	4,7	1,599314
Gómez, 2011 [158]	947,07	4,68	1,68623
Petcu et al., 2016 (Joback) [159]	4019,29	2,43	
Petcu et al., 2016 (Constantinou & Gani) [159]	977,88	3,33	
Petcu et al., 2016 (Gani & Marrero) [159]	1035,11	7,28	
Petcu et al., 2016 (Fedor) [159]	1043,28		
Philip, 2013 [160]	943,3	3,224	

**Table 3. Monoolein and Diolein critical properties. Parameters estimated by Valverde et al. [11] have followed the methods described in Reid, Prausnitz & Poling [161].**

Author	Monoolein			Diolein		
	Tc (K)	Pc (bar)	w	Tc (K)	Pc (bar)	w
Glisic & Skala, 2009 [156]	835,06	10,56	1,5324	920,2	5,05	1,7632
Aspen Plus v9, 2016	885	12,4	0,9918	1025	7,92	2,5810
Fornari et al., 2011 (Fedor) [162]	874,2	11,9		989,8	8,8	
Baroi, 2014 [163]	855,95	11,54	1,11			
Valverde et al. (Gomez-Thodos) [11]			1,1096			2,1174
Valverde et al. (Antoine) [11]			0,7573			1,7686
Valverde et al. (Lee-Kesler) [11]			0,8049			1,9529

**Table 4. Glycerol critical properties. Parameters estimated by Valverde et al. [11] have followed the methods described in Reid, Prausnitz & Poling [161].**

Author	Tc (K)	Pc (bar)	w
Costa et al., 2011 [164]	726,15	67	0,516
Aspen Plus v9, 2016	850	75	0,51269
Anikeev & Fan I, 2013 [165]	725	66,7	0,51
Picou, 2012 [166]	850,15	75,000765	0,5163
Valverde et al. (Ambrose) [11]	931,916368	288,710188	
Valverde et al. (Jobak) [11]	725,81559	66,966498	
Valverde et al. (Fedor) [11]	696,283279		
Valverde et al. (Gomez-Thodos) [11]			0,5849
Valverde et al. (Antoine) [11]			1,6856
Valverde et al. (Lee-Kesler) [11]			1,7881

**Table 5. Methyl-oleate critical properties.**

Author	Tc (K)	Pc (bar)	w
Glisic & Skala, 2009 [156]	721,02	11,03	1,0494
Aspen Plus v9, 2016	764	12,8	1,0494
Singh & Medina, 1989 [167]	611,67	13,9524525	0,9896
Ruan et al., 2014 (Jobak) [168]	774,2	11,223	0,9784
Abbas Karim et al., 2011 [169]	866,94	11,22	0,9835
Fornari et al., 2011 (Fedor) [162]	784,2	10,5	
An et al., 2013 (Ambrose) [170]	772	11,68	0,998
An et al., 2013 (Fedor) [170]	794		

**Table 6. Carbon dioxide critical properties.**

Author	Tc (K)	Pc (bar)	w
Feroiu et al. 2013 [171]	304,1	73,8	
Budisa et al. 2014 [172]	304	73,8	
Goos et al. 2010 [173]	304,1282	73,773	
Aspen Plus v9, 2016	304,21	73,83	0,223621
Barton, 1991 [151]			0,225
Perry, 2008 [152]	304,21	73,9	0,224

## On BIPs literature values and estimation

**Table 8. Peng-Robinson EOS with VdW mixing rules Binary Interaction Parameters (BIPs) for the binary system methanol-triolein, methanol-glycerol, FAME-CO<sub>2</sub> and FAME-methanol.**

Author	MeOH-Triolein		MeOH-Glycerol		FAME-CO <sub>2</sub>		FAME-MeOH	
	T (K)	k <sub>ij</sub>	T (K)	k <sub>ij</sub>	T (K)	k <sub>ij</sub>	T (K)	k <sub>ij</sub>
Tang et al., 2005 [174] (*)	353,2–463,2	0,0289						
Arevalos et al., 2013 [175]	473	-0,0657	493	0,0454	313,15	0,0397		
Arevalos et al., 2013 [175]	483	-0,12	523	0,0436	323,15	0,0409	523	0,0157
Arevalos et al., 2013 [175]	493	-0,298	543	0,0537	333,15	0,0431	548	0,0066
Arevalos et al., 2013 [175]	503	-0,25	573	0,0354	343,15	0,043	573	0,003

(\*) Unique BIP value obtained by considering  $l_{ij}$  zero.

**Table 9. Peng-Robinson EOS with VdW mixing rules Binary Interaction Parameters (BIPs) for the binary system methanol-carbon dioxide.**

Author	T (K)	$k_{ij}$
Lopez-Castillo et al., 2008 [176]	313,15	0,06
Hughes et al., 2014 [177]	213,15-247,12	0,0547
Chin et al., 2010 [178]	303,15	0,05506
Chin et al., 2010 [178]	313,15	0,068

**Table10. Peng-Robinson EOS with VdW mixing rules Binary Interaction Parameters (BIPs) for the binary system triolein-carbon dioxide.**

Author	T (K)	$k_{ij}$
You & Lee, 1993 [179]	313,15	0,082
You & Lee, 1993 [179]	323,15	0,083
You & Lee, 1993 [179]	333,15	0,084

**Table 11. Peng-Robinson EOS with VdW mixing rules Binary Interaction Parameters (BIPs) for the binary system glycerol-carbon dioxide.**

Author	T (K)	$k_{ij}$
Medina-Gonzalez et al., 2013 [180]	423,15	0,175175
Medina-Gonzalez et al., 2014 [180]	453,15	0,250175
Medina-Gonzalez et al., 2015 [180]	478,15	0,312675
Arevalos et al., 2013 [175]	313,15	0,0601
Arevalos et al., 2013 [175]	353,15	0,0896

**Table 12. BIPs equation coefficients fitted from literature values at the temperature range of operation in this work.**

$k_{ij}^n$	T (°C)	MeOH-Triolein		MeOH-Glycerol		MeOH-CO <sub>2</sub>		Triolein-CO <sub>2</sub>	
		$k_{ij}^n$	$k_{ij}$	$k_{ij}^n$	$k_{ij}$	$k_{ij}^n$	$k_{ij}$	$k_{ij}^n$	$k_{ij}$
$k_{ij}^1$	150	11,43	0,0493	3,1250	-0,0202	-0,2722	0,2086	0,0507	0,0930
$k_{ij}^2$	180	-0,0142	-0,0234	-0,0029	0,0176	0,0012	0,2459	0,0001	0,0960
$k_{ij}^3$	205	-2266	-0,1180	-805	0,0371	-9,7840	0,2768	0,0031	0,0985

$k_{ij}^n$	T (°C)	FAME-CO <sub>2</sub>		FAME-MeOH		Glycerol-CO <sub>2</sub>	
		$k_{ij}^n$	$k_{ij}$	$k_{ij}^n$	$k_{ij}$	$k_{ij}^n$	$k_{ij}^n$
$k_{ij}^1$	150	0,6848	0,0311	-2,4941	0,1263	-2,7870	0,1843
$k_{ij}^2$	180	-0,0009	0,0210	0,0022	0,0780	0,0045	0,2488
$k_{ij}^3$	205	-111,9	0,0109	722,5860	0,0485	451,2	0,3093

## On Optimization Discussion

Table 13. Comparison of the objective function value after BIPs optimization ( $OF_1$ ) between genetic algorithm (GA) and GlobalSearch algorithms.

T (°C)	1st iteration $OF_1$		2nd iteration $OF_1$		3rd iteration $OF_1$	
	GA	GlobalSearch	GA	GlobalSearch	GA	GlobalSearch
150	1,2212	1,2258	1,2247	1,2253	1,2248	1,2253
180	0,2348	0,2341	0,2339	0,2341	0,2339	0,2339
205	0,2922	0,2600	0,2484	0,2635	0,2417	0,2549

Table 14. Value of the objective function (OF) for each iteration of GlobalSearch algorithm for the determination of de kinetic constants and the unknown BIPs.  $OF_0$  corresponds to the objective function value resulted from kintecic parameters optimization while  $OF_1$  is related to that obtained from BIPs optimization.

1 <sup>st</sup> iteration		2 <sup>nd</sup> iteration		3 <sup>rd</sup> iteration	
$OF_0$	$OF_1$	$OF_0$	$OF_1$	$OF_0$	$OF_1$
1,2268	1,2258	1,2253	1,2253	1,2253	1,2253
0,2357	0,2341	0,2341	0,2341	0,2339	0,2339
0,2623	0,2600	0,2652	0,2635	0,2550	0,2549

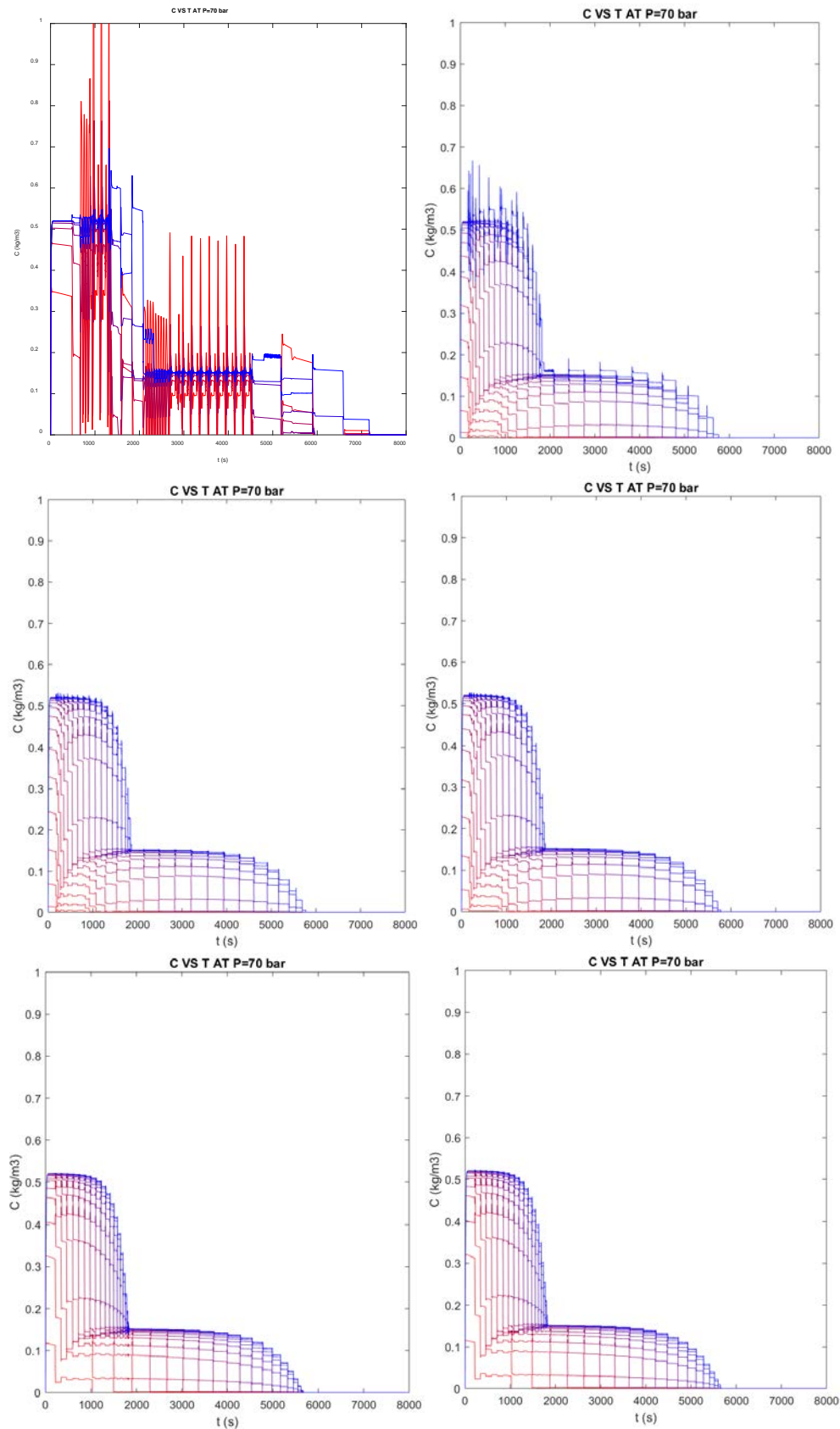


## Annex II

Table 15. Expressions of different orthogonal polynomials families.

Polynomial	Form	Expression of the formula	Domain
<b>Jacobi (*)</b>	Rodrigues Formula	$\tilde{P}_n^{\alpha,\beta} = \frac{(-1)^n}{n!} (1-x)^{-\alpha} x^{-\beta} \frac{d^n}{dx^n} (1-x)^{n+\alpha} x^{n+\beta}$	$x \in [0,1]$
	Series	$\tilde{P}_n^{\alpha,\beta} = (-1)^n \sum_{k=0}^{\min\{n,n+\alpha\}} \binom{n+\alpha}{k} \binom{n+\beta}{n-k} (1-x)^{n-k} (-x)^k$	
<b>Jacobi (*)</b> <b><math>\alpha=1</math></b> <b><math>\beta=0</math></b>	Rodrigues Formula	$\tilde{P}_n^{1,0} = \frac{(-1)^n}{n!} \frac{d^n}{dx^n} (1-x)^{n+1} x^n$	$x \in [0,1]$
	Series	$\tilde{P}_n^{1,0} = (-1)^n \sum_{k=0}^n \binom{n}{k}^2 \frac{n+1}{n-k+1} (1-x)^{n-k} (-x)^k$	
<b>Legendre (*)</b> <b>(Jacobi</b> <b><math>\alpha=0 \beta=0</math>)</b>	Rodrigues Formula	$\tilde{P}_n = \frac{(-1)^n}{n!} \frac{d^n}{dx^n} (1-x)^n x^n$	$x \in [0,1]$
	Series	$\tilde{P}_n = (-1)^n \sum_{k=0}^n \binom{n}{k} \binom{n+k}{k} (-x)^k$	
<b>Hermite</b>	Rodrigues Formula	$H_n = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$	$x \in [-\infty, \infty]$
	Series	$H_n = \begin{cases} (2x)^n \sum_{k=0}^{n/2} \frac{n! (-1)^k}{k! (n-2k)!} (2x)^{-2k} & \text{if } n \text{ pair} \\ (2x)^n \sum_{k=0}^{(n-1)/2} \frac{n! (-1)^k}{k! (n-2k)!} (2x)^{-2k} & \text{if } n \text{ odd} \end{cases}$	
<b>Laguerre</b>	Rodrigues Formula	$L_n^\alpha = x^{-\alpha} e^x \frac{d^n}{dx^n} x^{n+\alpha} e^{-x}$	$x \in [0, \infty]$
	Series	$L_n^\alpha = \sum_{k=0}^n (-1)^k \frac{n! (n+\alpha)!}{k! (n-k)! (\alpha+k)!} x^k$	
<b>Laguerre</b> <b><math>\alpha=0</math></b>	Rodrigues Formula	$L_n = e^x \frac{d^n}{dx^n} x^n e^{-x}$	$x \in [0, \infty]$
	Series	$L_n = \sum_{k=0}^n (-1)^k \binom{n}{k}^2 (n-k)! x^k$	
<b>Tchebycheff</b> <b>(Jacobi</b> <b><math>\alpha=-1/2</math></b> <b><math>\beta=-1/2</math>)</b>	Rodrigues Formula	$\tilde{T}_n = \frac{(-1)^n 2^{2n} n!}{(2n)!} \sqrt{x(1-x)} \frac{d^n}{dx^n} x^{n-1/2} (1-x)^{n-1/2}$	$x \in [0,1]$
	Series	$\tilde{T}_n = \begin{cases} \frac{n}{2} \sum_{k=0}^{n/2} (-1)^k \frac{(n-k-1)!}{k! (n-2k)!} (4x-2)^{n-2k} & \text{if } n \text{ pair} \\ \frac{n}{2} \sum_{k=0}^{(n-1)/2} (-1)^k \frac{(n-k-1)!}{k! (n-2k)!} (4x-2)^{n-2k} & \text{if } n \text{ odd} \end{cases}$	

(\*)  $(-1)^n$  factor in Jacobi and Legendre polynomials is not needed, as the resulting polynomials are also orthogonal. This variation is used by Villadsen and Stewart [100] and Finlayson [102].



**Fig 16. Resulting plot featuring the concentration of extracted lanolin in the solvent over time at 70 bar, 30 °C, and 4 kg/h flow rate, applying diverse polynomials for the orthogonal collocation method. From up to down and left to right: Laguerre ( $\alpha=0$ ) and Hermite (same plot), Jacobi ( $\alpha=1, \beta=0$ ), Legendre, Tchebycheff, Legendre (even polynomials), Jacobi ( $\alpha=1, \beta=0$ , even polynomials).**



## Annex III

### On Inlet Breakthrough Curve's Demonstration

$$-\frac{1}{V} \frac{\partial m_L}{\partial t} = k_G a (C_g^* - C_g) \quad (1)$$

Eqn. (1) is a simple mass balance considering that the lanolin concentration gradient over time must be equal to the concentration gradient in the solvent phase. From now on, four assumptions are made:

- 1) as lanolin layer is thin, the mass transfer area  $a$  is assumed constant,
- 2) the radial distribution of lanolin mass in the liquid layer  $m_L$  in a certain moment follows the same profile as the evolution of the lanolin concentration in the liquid-solvent interphase over time (similar to an ergodic argument),
- 3) the lanolin concentration in the liquid phase straight over the inner wool fiber, namely at the inner radius  $R$ , can be assumed constant and equal to  $C_{L0}$  (the lanolin density),
- 4) equilibrium of the lanolin concentration at the liquid side and at the solvent side of the liquid-solid interphase can be modeled by a Henry type equilibrium  $C_g^* = K C_L^*$ , where  $K$  is the Henry constant for the lanolin-solvent binary system.

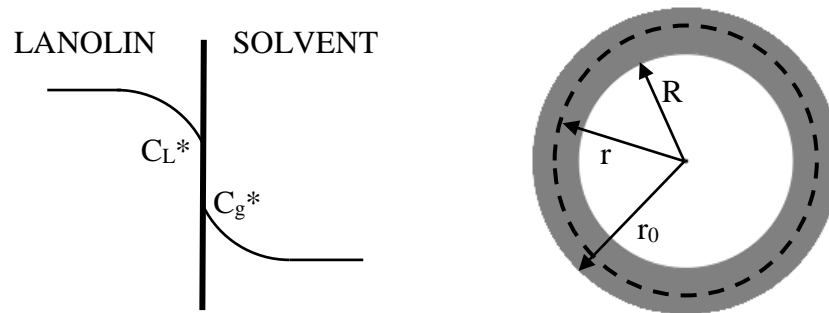
As it can be seen, those assumptions avoid considering a constant radial profile of lanolin concentration in the liquid phase, approaching to a more realistic model.

On the other hand, dynamic equilibrium between liquid diffusion and solvent mass transfer at the liquid-solvent interphase must be satisfied.

$$-D \frac{\partial C_L^*}{\partial r} = k_g (K C_L^* - C_g) \quad (2)$$

If radial profile of lanolin concentration in the liquid phase is unknown, as it is for the complexity of the diverse lanolin solubilities inside this layer, it is hard to solve considering both  $C_g^*$  and  $C_g$ . Hereby, this model will be found only at the extractor inlet, were initial condition  $C_g(z,t) = 0$  simplifies both Eqn. (1) and (2).

Where  $k_g$  is the mass transfer coefficient in the solvent,  $D$  is the molecular diffusivity in the liquid phase,  $C_L^*$  is the lanolin concentration at the liquid side of the interphase.



**Fig 1. Liquid-solvent interphase profile and wool fiber section**

Integrating Eqn. (2) from  $r$  to the inner wool fiber radius  $R$ .

$$\ln\left(\frac{C_L^*}{C_{LR}^*}\right) = \frac{k_g K}{D} (R - r) \quad (3)$$

Where  $C_{LR}^*$  would be the lanolin concentration in the liquid phase when the interphase was at the inner wool fiber radius  $R$ , namely when all lanolin would be extracted. As said, it is assumed to be constant and equal to  $C_{L0}$ . Now, considering the second assumption, Eqn. (4) make sense.

$$\frac{\partial m_L}{\partial r} = C_L^* 2\pi r L N_f \quad (4)$$

Where  $L$  is the fiber length and  $N_f$  the number of fibers present inside the extractor. Here, radial lanolin density is equated to the lanolin concentration at the liquid side of the interphase multiplied by the interphase surface, considering  $r$  as the interphase radius.

Now, isolating  $C_L^*$  from Eqn. (3), replacing in Eqn. (4) and integrating:

$$m_L = \int_r^{r_0} C_L^* 2\pi r L N_f dr = C_{LR}^* 2\pi L N_f \frac{D}{k_g K} \left( \left( r + \frac{D}{k_g K} \right) e^{\frac{k_g K}{D}(R-r)} + C e^{\frac{k_g K}{D}R} \right) \quad (5)$$

Where  $C$  is the integration constant, which can be obtained from the initial condition, namely when the interphase radius is  $r_0$  the lanolin mass must be equal to the initial total one  $m_{L0}$ .

$$m_L = m_{L0} + C_{LR}^* 2\pi L N_f \frac{D}{k_g K} \left( \left( r + \frac{D}{k_g K} \right) e^{\frac{k_g K}{D}(R-r)} - \left( r_0 + \frac{D}{k_g K} \right) e^{\frac{k_g K}{D}(R-r_0)} \right) \quad (6)$$

Deriving with respect to time:

$$\frac{\partial m_L}{\partial t} = -C_{LR}^* 2\pi L N_f r e^{\frac{k_g K}{D}(R-r)} \frac{\partial r}{\partial t} \quad (7)$$

Where  $\partial r / \partial t$  is obtained from deriving Eqn. (3).

$$\frac{\partial r}{\partial t} = -\frac{D}{k_g K} \frac{1}{C_L^*} \frac{\partial C_L^*}{\partial t} \quad (8)$$

In Eqn. (7) radius  $r$  is replaced by the expression resulted from isolating it in the Eqn. (3), and  $\partial r / \partial t$  is replaced by Eqn. (8).

$$\frac{\partial m_L}{\partial t} = -2\pi L N_f \frac{D}{k_g K} \left( \frac{D}{k_g K} \ln \left( \frac{C_L^*}{C_{LR}^*} \right) - R \right) \frac{\partial C_L^*}{\partial t} \quad (9)$$

Replacing Eqn. (9) in Eqn. (1) (remember  $C_g$  is equal to zero at the extractor inlet):

$$\frac{1}{V} 2\pi L N_f \frac{D}{k_g K} \left( \frac{D}{k_g K} \ln \left( \frac{C_L^*}{C_{LR}^*} \right) - R \right) \frac{\partial C_L^*}{\partial t} = k_g a K C_L^* \quad (10)$$

Considering the number of wool fibers present in the extractor as:

$$N_f = \frac{m_{wool_0}}{\pi r_0^2 L \rho_{wool}} = \frac{V_{wool_0}}{\pi r_0^2 L} \rightarrow \pi L N_f = \frac{V_{wool_0}}{r_0^2}$$

$$\frac{1}{V} 2\pi L N_f = \frac{V_{wool_0}}{V} \frac{2}{r_0^2} = \frac{2(1-\epsilon)}{r_0^2} \quad (11)$$

Now replacing in (10):

$$\frac{2(1-\epsilon)}{r_0^2} \frac{D}{k_g K} \left( \frac{D}{k_g K} \ln \left( \frac{C_L^*}{C_{LR}^*} \right) - R \right) \frac{\partial C_L^*}{\partial t} = k_g a K C_L^* \quad (12)$$

Solving differential equation (12) by separating variables and integrating from  $t=0$  to any time  $t$ :

$$\int_{C_{L_0}^*}^{C_L^*} \left( \frac{D}{k_g K} \frac{\ln \left( \frac{C_L^*}{C_{LR}^*} \right)}{C_L^*} - \frac{R}{C_L^*} \right) \partial C_L^* = \int_0^t \frac{r_0^2 (k_g a K)^2 a}{2D(1-\epsilon)} \partial t$$

$$\frac{D}{2k_g K} \left( \ln^2 \left( \frac{C_L^*}{C_{LR}^*} \right) - \ln^2 \left( \frac{C_{L_0}^*}{C_{LR}^*} \right) \right) - R \ln \left( \frac{C_L^*}{C_{L_0}^*} \right) = \frac{r_0^2 (k_g a K)^2 a}{2D(1-\epsilon)} t \quad (13)$$

Isolating and replacing  $C_{LR}^* = C_{L_0} = \rho_{LAN}$ :

$$\ln^2 \left( \frac{C_L^*}{C_{L_0}^*} \right) - \frac{2Rk_g K}{D} \ln \left( \frac{C_L^*}{C_{L_0}^*} \right) - \frac{r_0^2 (k_g a K)^3 a}{D^2(1-\epsilon)} t = 0 \quad (14)$$

Replacing  $x = \ln \left( \frac{C_L^*}{C_{L_0}^*} \right)$  a second-degree equation is obtained, from which negative solution is the correct profile:

$$C_L^* = C_{L_0}^* e^{\frac{Rk_g K}{D} \left( 1 - \sqrt{1 + \frac{r_0^2 k_g a K t}{R^2 (1-\epsilon)}} \right)} \quad (15)$$

As discussed in Valverde et al. [62], Eqn. (15) can be expressed in terms of Biot number and a certain A constant, leading to a Weibull type expression.

$$C_L^* = C_{L_0}^* e^{Bi} e^{-Bi\sqrt{1+At}} = C_0 e^{-\alpha(n+t)^b} \quad (16)$$

Were  $b$  in the extractor inlet case is  $1/2$ ,  $n$  is the extraction time delay since solvent flowrate enters the extractor until it exits,  $\alpha$  is related to the Biot number and  $C_0$  is related to the maximum concentration of lanolin in the liquid. Biot number and A parameters are defined as follow:

$$Bi = \frac{Rk_g K}{D} \quad (17)$$

$$A = \frac{r_0^2}{R^2} \frac{k_g a K}{1-\epsilon} \quad (18)$$

Replacing Eqn. (15) in the extraction rate expression, namely the left side of Eqn.

(1):

$$r_v = k_g a K C_{L_0}^* e^{\frac{Rk_g K}{D} \left( 1 - \sqrt{1 + \frac{r_0^2 k_g a K t}{R^2 (1-\epsilon)}} \right)} \quad (19)$$

Matching Eqn. (19) with the balance equation for dispersed flow:

$$\epsilon \frac{\partial C_g}{\partial t} + u \frac{\partial C_g}{\partial z} = k_g a K C_{L_0}^* e^{\frac{Rk_g K}{D} \left( 1 - \sqrt{1 + \frac{r_0^2 k_g a K t}{R^2 (1-\epsilon)}} \right)} \quad (20)$$

Eqn. (20) expression allows finding the breakthrough curve at a differential step of the extractor inlet, and it relates the lanolin concentration in the solvent,  $C_g$ , with the Weibull distribution seen in Eqn. (16). Hence, although parameters found for  $C_L^*$  at the extractor inlet will be different, it can be assumed that the same Weibull type function can model the breakthrough curve at the extractor outlet:

$$C_g(Z_T, t) = c' e^{-\alpha(n+t)^b} \quad (21)$$

Nevertheless, experimental values from Eychenne et al. [66] are lanolin extracted fractions. Thus:

$$X(t) = \frac{\int_0^t c_g(t, z_T) dt}{\int_0^\infty c_g(t, z_T) dt} \quad (22)$$

$$X(t) = c \int_0^t e^{-\alpha(t+n)^b} dt \quad (23)$$

To simplify the fitting of parameters inside an integral expression, an analytical solution of the integral was found, starting by solving the case without time delay  $n$ .

$$X(t) = c \int_0^t e^{-\alpha t^b} dt \quad (24)$$

Replacing variables:

$$\alpha = \omega^b \quad \text{and} \quad x^{1/b} = \omega t \quad (25)$$

$$dt = \frac{1}{b\omega} x^{\frac{1}{b}-1} dx \quad (26)$$

$$X(x) = \frac{c}{b\omega} \int_0^x x^{\frac{1}{b}-1} e^{-x} dx = \frac{c}{b\omega} \gamma\left(\frac{1}{b}, x\right) \quad (27)$$

Where  $\gamma(\frac{1}{b}, x)$  is the incomplete Euler's gamma function with  $1/b$  argument, which might not to be integer. This function can be approximated by Taylor power series, such as those by Abramowitz i Stegun [181]:

$$\gamma(m, x) = x^m \sum_{k=0}^{\infty} \frac{(-x)^k}{k!(m+k)} \text{ si } m > 0 \quad (28)$$

Or by Press et al. [182]:

$$\gamma(m, x) = e^{-x} x^m \sum_{k=0}^{\infty} \frac{x^k}{(m+k)(m+k-1)\dots m} \quad (29)$$

The latter is very suitable for  $x \ll m$  and for arguments not necessarily integers nor positive. In this case, if time is expressed in hours, with enough terms of the series from Eqn. (29) a good approximation with little error can be found. Thus, applying Eqn. (29) to Eqn. (27):

$$X(x) = \frac{c}{b\omega} e^{-x} x^{\frac{1}{b}} \sum_{k=0}^{\infty} \frac{x^k}{(\frac{1}{b}+k)(\frac{1}{b}+k-1)\dots\frac{1}{b}} \quad (30)$$

Replacing original variables  $x = \omega^b t^b = \alpha t^b$ :

$$X(t) = \frac{c}{b\alpha^{\frac{1}{b}}} e^{-\alpha t^b} \alpha^{\frac{1}{b}} t \sum_{k=0}^{\infty} \frac{(\alpha t^b)^k}{(\frac{1}{b}+k)(\frac{1}{b}+k-1)\dots\frac{1}{b}} \quad (31)$$

$$X(t) = ce^{-\alpha t^b} t \left( 1 + \frac{\alpha t^b}{\frac{1}{b}+1} + \frac{(\alpha t^b)^2}{(\frac{1}{b}+1)(\frac{1}{b}+2)} + \frac{(\alpha t^b)^3}{(\frac{1}{b}+1)(\frac{1}{b}+2)(\frac{1}{b}+3)} + \dots \right) \quad (32)$$

Applying the temporal translation:

$$X(t) = ce^{-\alpha(t-n)^b} (t-n) \left( 1 + \frac{\alpha(t-n)^b}{\frac{1}{b}+1} + \frac{(\alpha(t-n)^b)^2}{(\frac{1}{b}+1)(\frac{1}{b}+2)} + \frac{(\alpha(t-n)^b)^3}{(\frac{1}{b}+1)(\frac{1}{b}+2)(\frac{1}{b}+3)} + \dots \right) \quad (33)$$

The limit of Eqn. (33) can be found exactly through Eqn. (27).

$$\lim_{x \rightarrow \infty} \frac{c}{b\omega} \gamma\left(\frac{1}{b}, x\right) = \frac{c}{b\omega} \Gamma\left(\frac{1}{b}\right) = \frac{c}{b\alpha^{\frac{1}{b}}} \Gamma\left(\frac{1}{b}\right) \quad (34)$$

If it is imposed that this limit is equal to one, as obviously the extracted fraction should tend to one at infinite time, expression will depend on one parameter left:

$$\frac{c}{b\alpha^{\frac{1}{b}}} \Gamma\left(\frac{1}{b}\right) = 1 \rightarrow \alpha = \left(\frac{c}{b} \Gamma\left(\frac{1}{b}\right)\right)^b \quad (35)$$

Replacing in Eqn. (32):

$$X(t) = ce^{-\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)t\right)^b} t \left( 1 + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)t\right)^b}{\frac{1}{b}+1} + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)t\right)^{2b}}{\left(\frac{1}{b}+1\right)\left(\frac{1}{b}+2\right)} + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)t\right)^{3b}}{\left(\frac{1}{b}+1\right)\left(\frac{1}{b}+2\right)\left(\frac{1}{b}+3\right)} + \dots \right) \quad (36)$$

With the temporal translation:

$$X(t) = ce^{-\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^b} (t-n) \left( 1 + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^b}{\frac{1}{b}+1} + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^{2b}}{\left(\frac{1}{b}+1\right)\left(\frac{1}{b}+2\right)} + \frac{\left(\frac{c}{b}\Gamma\left(\frac{1}{b}\right)(t-n)\right)^{3b}}{\left(\frac{1}{b}+1\right)\left(\frac{1}{b}+2\right)\left(\frac{1}{b}+3\right)} + \dots \right) \quad (37)$$

This last expression depending on three parameters is the one that best represents the behavior of the lanolin extracted fraction curve over time. With a simple fitting tool from a mathematical software like MATLAB, it is possible to find both extracted fraction curve and concentration breakthrough curve of an extraction process.

## On Weibull Distribution Results

**Table 17. Parameters of fitting Weibull's cumulative distribution function integral to experimental data. R-squared ( $R^2$ ) is shown as a measure of the goodness of fit.  $\alpha$  has been calculated applying Eqn. (35) after finding  $c$  and  $b$ .  $n$  is the  $n$  parameter used in the fitting, and  $n'$  the theoretical value that should be (the same as  $n$  if empty).  $c'$  is the solubility value obtained from Weibull distribution while  $sol.$  is the solubility extracted from Eychenne et al. [66].**



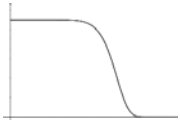
P bar	T °C	Q kg/h	$\rho_B$ kg/m <sup>3</sup>	R <sup>2</sup>	c	$\alpha$	b	n (*)	n'	n-n'	c' (**) kg/m <sup>3</sup>	sol. kg/ m <sup>3</sup>
120	80	3	127	0,999	0,183	0,560	1,023	0,0111			<b>0,057</b>	0,33
120	80	3	227	0,999	0,174	0,510	1,216	0,0057			<b>0,054</b>	0,33
120	80	4	159	0,999	1,430	1,524	1,738	0,1415	0,0065	0,135	<b>0,33</b>	0,33
120	80	4	227	0,995	1,227	1,271	2,735	0,1305	0,0043	0,126	<b>0,29</b>	0,33
120	80	4	318	0,997	1,599	1,837	1,713	0,1346	0,0028	0,131	<b>0,37</b>	0,33
120	80	5	127	0,988	2,194	2,279	1,097	0,1350	0,0092	0,126	<b>0,41</b>	0,33
120	80	5	227	0,999	1,983	2,175	1,276	0,0051			<b>0,37</b>	0,33
150	60	4	127	0,998	1,669	1,772	1,340	0,0155			<b>0,73</b>	0,74
150	80	4	127	0,997	2,114	2,338	1,255	0,0115			<b>0,68</b>	0,69
200	60	4	127	0,991	2,556	3,073	1,309	0,0176			<b>1,26</b>	1,24

(\*)  $n = Z_T \epsilon / u = Z_T \pi R_B^2 \epsilon \rho_s / Q$ , where  $Z_T$  is the extractor height,  $\epsilon$  the bed porosity,  $u$  the superficial velocity,  $R_B$  the bed radius,  $\rho_s$  the solvent density and  $Q$  the mass flowrate.

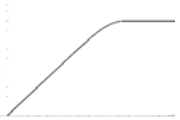
(\*\*)  $c' = c_M \rho_s / Q$  as  $C_g(Z_T, t) = (dX/dt) m_L \rho_s / Q$ , where  $m_L$  is the mass of lanolin,  $C_g(Z_T, t)$  the concentration of lanolin in the solvent at the extractor's outlet and  $X$  the extracted fraction of lanolin.

## On Solving Method Abstract

**EXPERIMENTAL DATA AUGMENTATION**

$C_g = c' e^{-a(t+n)^b}$ 


Fit experimental extraction curves  $X(t)$  with the integral of the Weibull's distribution function.


 $X = \int c e^{-a(t+n)^b} dt$

**DEFINITION OF THE NETWORK**

Create a recurrent network (*NARXNET*) with two different kind of inputs: a cell with the variables  $P, T, Q$  and  $\rho_B$  to define the extractor's state of operation, called the external inputs; and a cell with the extraction data orded in time.

$$U = \begin{Bmatrix} P_i \\ T_i \\ Q_i \\ \rho_{B_i} \end{Bmatrix}$$

→

U TIME

→

FEED FORWARD NET

→

$$\{X_i^t\}$$

$$Y = \{X_i^{t-1}\}$$

→

Y TIME

→

FEED FORWARD NET

→

$$\{X_i^t\}$$

**NETWORK TRAINING**

The network must learn recurrently in time with the state variables and the extraction values of the previous time. Thus, the external input consists of a cell array with as many cells as time steps. In each cell, an array with the  $[P, T, Q, \rho_B]$  variables of each different status is provided. The feedback input will consist of a cell array of the same length, in which in every cell an array with the extracted fractions at each time and state will be provided. The time delay will be 1.

$\begin{Bmatrix} P_i & P_{i+1} \\ T_i & T_{i+1} \\ Q_i & Q_{i+1} \\ \rho_{B_i} & \rho_{B_{i+1}} \end{Bmatrix}$	$\begin{Bmatrix} P_i & P_{i+1} \\ T_i & T_{i+1} \\ Q_i & Q_{i+1} \\ \rho_{B_i} & \rho_{B_{i+1}} \end{Bmatrix}$	...
--	--	-----

$\{X_i^{t_0}, X_{i+1}^{t_0}, \dots\}$	$\{X_i^{t_1}, X_{i+1}^{t_1}, \dots\}$	...
---------------------------------------	---------------------------------------	-----

U TIME DELAY = 1	Y TIME DELAY = 1
---------------------	---------------------

**NETWORK TESTING**

One experimental data set at a  $[P_j, T_j, Q_j, \rho_{B_j}]$  state was not included in training. The extraction curve for this state is obtained introducing the array of the state as an external input and the first feedback value  $X_j^{t_0} = 0$ . The NN is used in closed loop as many time steps ahead as needed. Error of prediction is mesured with AARD (Average Absolute Relative Deviations).

