Review

Journal of the Brazilian Chemical Society

https://dx.doi.org/10.21577/0103-5053.20250084

J. Braz. Chem. Soc. **2025**, *36*, 8, e-20250084, 1-17

Determination of Thermal Process Kinetics by Artificial Neural Networks: A Review Considering Isothermal and Non-Isothermal TG and DSC Data

Rita C. O. Sebastião,[*,a] *Natália R. S. Araujo,*[a] *Felipe S. Carvalho,*[b] *Bárbara D. L. Ferreira*[a] *and João Pedro Braga*[a]

[a]*Laboratório de Problemas Inversos e Cinética Química (PINCQ), Departamento de Química, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte-MG, Brazil*

[b]*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena 91125, California, USA*

The integration of artificial neural networks (ANNs) with thermal analysis techniques, such as thermogravimetry (TG) and differential scanning calorimetry (DSC), presents great potential to accurately determine kinetics of thermal processes and thermal behavior. This review examines the application of ANNs, including Multilayer Perceptron (MLP) and Hopfield Neural networks (HNN), to analyze both isothermal and non-isothermal TG and DSC data. These networks provide robust tools for predicting thermal behavior, determining kinetic parameters, and modeling complex reaction mechanisms, significantly improving accuracy when compared to traditional methods. The review highlights advances in ANN-based methodologies, including the ability to integrate multiple kinetic models and manage noisy experimental data, thereby offering new insights into thermal decomposition and reaction kinetics. Moreover, the use of HNN in distributed activation energy models (DAEM) is explored, indicating their potential in understanding heterogeneous systems. This study emphasizes the transformative potential of ANNs for improving kinetic studies, suggesting future directions to explore machine learning in material characterization and thermal process optimization.

Keywords: artificial neural networks, thermal analysis, thermogravimetry, differential scanning calorimetry, kinetics, machine learning

1. Introduction

Nowadays, the world is experiencing a new industrial revolution characterized by Artificial Intelligence's (AI) development and improvement.[1,2] Among the various branches of AI, machine learning (ML) stands out, since it can identify patterns in data sets to predict results without relying on pre-programmed instructions.[3]

Artificial Neural Networks (ANNs) are a type of machine learning algorithm modeled to resemble the biological neural networks found in the brain.[4] A specific type, called Multilayer Perceptron Neural (MLP) Network, presents a structure consisting of interconnected nodes or artificial neurons arranged in layers that can be trained to perform various tasks. Each node in an ANN receives one or more inputs, performs a calculation, and passes the result on to the next layer of nodes. Nodes in the first layer receive input data, while nodes in the last layer produce the final output of the network.

ANNs have been successfully applied to pattern recognition,[5] image and speech recognition,[6,7] natural language processing,[8] automatic control,[9] financial forecasting[10] among several other applications. They have revolutionized many fields by providing new ways to analyze and understand complex data, and they continue to be an active area of research in machine learning and artificial intelligence.

Thermal analysis is a set of techniques that aims to measure physical properties of a substance and/or its reaction products as a function of time or temperature, when this sample is subjected to a controlled temperature program and under a specified atmosphere.[11] This paper will focus on discussing the application of thermogravimetry (TG)

*e-mail: ritacos@ufmg.br
Editor handled this article: Maurício Coutinho-Neto (Guest)

and differential scanning calorimetry (DSC), techniques in which the measured physical phenomenon is the variation of mass and heat flux of the sample, respectively, as a function of temperature and/or time.

ANNs can be categorized into three types, based on interpretability: white-box, gray-box, and black-box models.[12,13] White-box models use explicit physical equations. In thermal analysis, an example is the use of kinetic theories applied to extract parameters, such as the triplet kinetic parameters (activation energy, E_a, pre-exponential factor, A, and reaction model), from DSC or TGA data. These models are fully interpretable but struggle with complex, multi-step reactions.[14] Examples of this approach include traditional isoconversional or model-fitting methods.[15,16] Black-box models are purely data-driven, in which the ANNs model learns input-output patterns without using physical laws.[17,18] They handle noise and nonlinearity well but lack mechanistic insight and may fail outside their training range.[14]

The gray-box model combines both white-box and black-box models. This network architecture integrates traditional modeling with ANNs by optimizing an unknown aspect of the process.[2] This type of model aims to achieve a balance between interpretability and flexibility but requires high reliability in the parameters and datasets during the training step.

In the last decades, studies[19-23] have proposed the application of neural networks to determine the values of activation energy, pre-exponential factor, and mechanism. In thermal decomposition processes, the mechanisms are generally complex and not well understood. In addition, to accurately describe the overall process, it is required to assume a combination of several models, constructing the mechanism as a combination of competing, independent, or consecutive steps.[24,25]

Determining kinetic parameters from experimental data is considered an ill-conditioned inverse problem, which means it may have more than one possible solution.[26] The decision about the right (or coherent) solution must be taken together with experimental researchers. It is not an exclusive mathematical issue, from which the best solution is selected by only considering the minimal residual error.

In contrast to thermodynamic quantities, no reference value for E_a can be found in the literature. Consequently, an ANN trained using wrong parameters (or patterns) may fail to learn and correlate correct values in patterns, even if it achieves low adjustment errors during the training step. This step of verifying the parameters can be executed using external methodologies, such as isoconversional analysis for non-isothermal data and Arrhenius analysis for isothermal data.[27-29] Also, the experimental conditions must be carefully considered in the ANN thermal behavior prediction,[30] as certain materials can present a significant change in the kinetic mechanism depending on the isothermal or non-isothermal process.

Works as the one proposed by Conesa *et al.*[20] simulate data with known kinetic parameters to train the neural network model. This approach is intended to overcome the lack of reference values for E_a. However, there are some points one should be aware of. Taking for example some of the values these authors[20] have used for generating the data: the pre-exponential factor $k_{ref} = 5 \times 10^{12}$ and $E_a/R = 10000$. These values yield the product $5e^{-10000}$. However, the same factor in the differential equation to be solved could be achieved by selecting $k_{ref} = 6 \times 10^{12}$ and $E_a/R = 10000 + \ln(6/5)$. Considering that $\ln(6/5)$ ca. 0.18, the change in exponent would be minimal. Thus, one would have certainty only about the rate constant and the model predictions may be biased by the parameter selection. This raises the question: what parameters should be used for simulating the training data? Should all possible combinations be used? If so, how to evaluate the predictions made by the ANN model?

This paper will cover the use of artificial neural networks in two techniques of thermal analysis: TG and DSC. The first works using ANNs were Sbirrazzuoli *et al.*[23,31] for DSC data, Sebastião *et al.*[19] for isothermal thermogravimetry and Conesa *et al.*[20] for thermogravimetry dynamic data. The objective is to review the leading applications in the field over time and the improvement due to the use of AI for deepening the knowledge in thermal analysis.

2. Methodology

2.1. Machine learning

For the successful development of machine learning, it is necessary to have a solid knowledge about the addressed problem and use reliable experimental data, since it is necessary to interpret the results acquired from a good experimental adjustment. These algorithms can be developed in five general steps, as shown in Figure 1.

The choice of the most appropriate machine learning algorithm depends on different factors such as, the physical characteristic of the problem, the number of variables, the mathematical model that best fit the experimental data, among others.[32] In thermal analysis, neural networks are commonly applied to investigate thermal behavior, thermal stability and the kinetics of the thermal processes, whether to simulate experimental data or to recover properties.

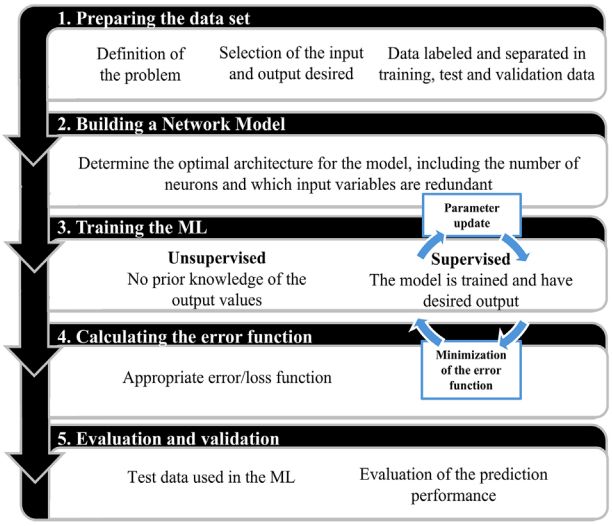

Figure 1. General steps of ML development.

2.2. Neural networks

The artificial neural network (ANN) is a powerful mathematical tool based on the biological communication model, used to solve linear and non-linear problems in various areas of science and has the artificial neuron as its processing unit.[33] These units are arranged in parallel layers, in one or more layers (multilayer perceptron neural network - MLP) or fully interconnected in a recurrent neural network (Hopfield neural network). Neurons communicate unidirectionally through synaptic connections, with their associated weights being adjusted systematically using the steepest descent algorithm, which updates the weights in the direction opposite to the gradient of the error function. The weights optimization process characterizes the training step of the usual neural networks.

Figure 2 shows a comparison between an artificial neuron and a biological neuron. In a biological neuron, dendrites receive signals from other neurons through synapses. The incoming signal is processed in the soma (cell body) and then sent through the axon, which forms a synapse with the next neuron. In the artificial model, the input layer values act as the external signal, while the synapse is represented by the weights. The processing of this signal, i.e., the summation of the weighted values and the application of an activation function, *f*, represents the processing that occurs in the cell body. Finally, this processed signal is transmitted to the next set of neurons through synapses, characterized by the next set of weights in the ANN architecture.

The ANN model was first proposed in 1943 by the neurophysiologist Warren Sturgis McCulloch and the mathematician Walter Harry Pitts Junior.[34] In their model, the input data, x_m, together with the synaptic weights w_{km}

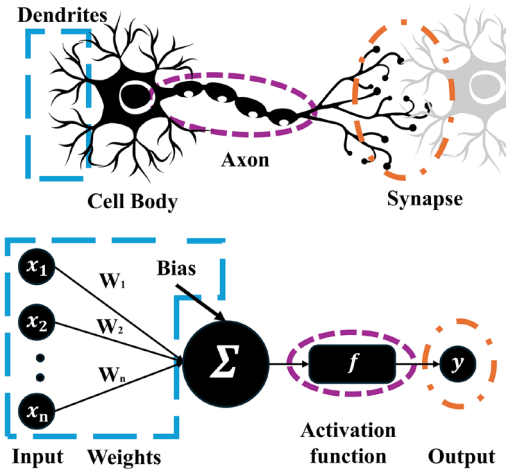

Figure 2. Comparison between artificial and biological neurons (adapted from reference 4).

are linearly combined by the additive function, equation 1, also considering the bias, b_k. The state of the neurons, u_k, is processed by the activation function *f*, generating the output y_k, equation 2.[4]

$$u_k = \left(\sum_{j=0}^{m} w_{kj} x_j \right) \quad (1)$$

$$y_k = f(u_k + b_k) \quad (2)$$

The term b_k in equation 2 is the bias, a set of positive or negative external parameters specific to each neuron, designed to increase or decrease the effect of the activation function. The activation function *f* in equation 2 can assume several mathematical functions depending on the problem. The activation functions used can be monotonically increasing, such as sigmoid or hyperbolic tangent.[35]

From the literature, in thermal analysis it is observed that Multilayer Perceptron (MLP) network is the most applied to isothermal and non-isothermal TG data and the Hopfield network (HNN) in DSC data. Thus, these two ANN will be discussed in the next sections.

2.3. Multilayer Perceptron (MLP) Neural Network

The MLP is a type of network containing successive layers of neurons, enabling it to efficiently tackle complex problems by learning hierarchical representations. Figure 3 schematizes an MLP network with three consecutive layers: the input, one hidden and the output layer.

The input data, which can be either experimental data or parameters, is fed into the input layer, as shown in Figure 3, which is commonly activated by a linear function. The information is then transferred to the neurons in the second layer, which are generally activated by a nonlinear activation

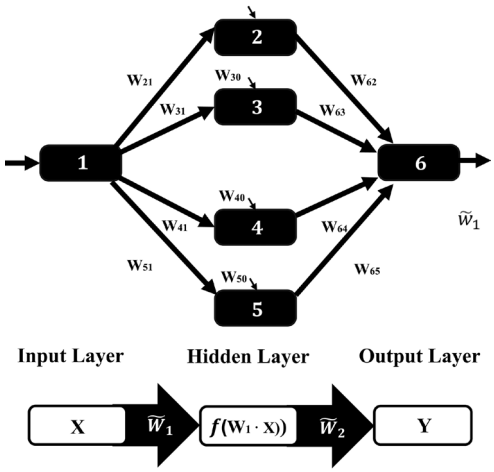

Figure 3. General Multilayer Perceptron Neural Network schematic architecture (adapted from reference 36).

function, such as tan $h(x)$ or *sigmoidal* = 1/1 + exp(–xt), depending on the problem. This layer is called hidden or intermediate layer and contains a single or multiple set of neurons. The response of the MLP is determined by the state of the neurons in the output layer, usually calculated as the linear combination of the neuron signals from the previous layer.

The backpropagation algorithm is widely used for supervised training of MLP networks and was proposed in 1986 by Rumelhart and Chauvin[37,38] This method is based on the Widrow-Hoff rule or delta rule, which minimizes the error in the output layer by updating the weights of the connections of the previous layers. For this reason, the activation functions used must be continuous and differentiable. The error function, E, to be minimized and the change of network connection weights, Δw_{ji}, are shown in equations 3 and 4, respectively,[4]

$$E = \frac{1}{2}\sum_{j=0}^{k}\left(d_i - y_i\right)^2 \tag{3}$$

$$\Delta w_{ji} = -\eta\,\frac{\partial E}{\partial w_{ji}} \tag{4}$$

where k is the number of output neurons in the network, d_i is the desired output and y_i is the output calculated by the network at time t, Δw_{ji} is the weight change of the connection between neurons i and j, η is the learning rate. When the error function shown in equation 3 represents a minimum value, this will be the acceptable network response.

2.4. Hopfield Neural Network

A Hopfield neural network (HNN) is a recurrent single-layer network where all logic units are interconnected. The neurons are connected by a weight factor, *Tij*,[39] between neurons i and j. The neuron state, u_i, is determined by the weighted sum of all neurons connected to it, as well as by external impulses, I_i, represented in equation 5. Figure 4 schematizes a Hopfield network.[40]

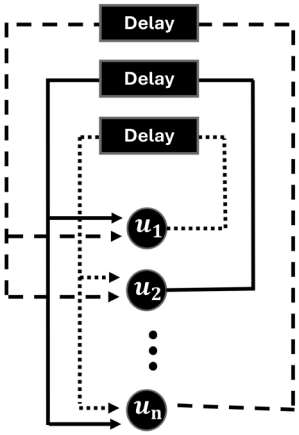

Figure 4. General Hopfield Neural Network Architecture (adapted from reference 4).

The neuron state is calculated by the contributions of all other neurons connected to it through a weighted sum of their inputs, as shown in equation 5. However, this contribution is only propagated if the activation function f_j is applied. During the learning time, τ, the information provided to the network is propagated and to maintain the stability of the network and avoid positive feedback, the diagonal elements of the symmetric matrix T_{ij} are null.[41]

$$\mu_i \frac{du_i(\tau)}{d\tau} = -u_i(\tau) + \sum_{j=1}^{n}\left[T_{ij}\,\mathrm{f_j}\left(\mathrm{u_j}(\tau)\right) + \mathrm{I_i}(\tau)\right] \tag{5}$$

In which the factor $I_i(\tau)$ represents an external stimulus to the neuron and $f_j(u_j(\tau))$ represents the activated state of the neuron. The energy function associated with each state of the network is given by equation 6.

$$E = \frac{1}{2}\sum_{j=1}^{m} e_j^2 = \frac{1}{2}\sum_{j=1}^{m}\left(C_{calc,j} - C_{exp,j}\right)^2 \tag{6}$$

In which e_j is the difference between $C_{calc,j}$, the calculated property and $C_{exp,j}$ is the experimental property, m is the number of experimental data. A set of differential equations, similar to the Hopfield differential equation 5,[42] is established by imposing $\frac{dE}{d\tau} < 0$. The error function expressed by equation 6 can be developed as,[43]

$$\frac{dE}{d\tau} = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(e_j\,\frac{\partial\left(C_{calc}\right)_j}{\partial f_i}\,\frac{\partial f_i}{\partial u_i}\,\frac{\partial u_i}{\partial \tau}\right) = \vec{\nabla}E.\frac{\vec{df}}{d\tau} < 0 \tag{7}$$

Where time-dependence of the neuron state is highlighted. Since $\frac{\partial f}{\partial u} > 0$ the condition of minimizing the error function is satisfied if $\frac{du_i}{d\tau} = -\frac{\partial E}{\partial f}$ or[26]

$$\frac{du_i}{d\tau} = -\sum_{j=1}^{m}\left(\frac{\partial\left(C_{calc}\right)_j}{\partial f_i}e_j\right) \quad (8)$$

Which validates the minimization of the error function because

$$\frac{dE}{d\tau} = -\sum_{i=1}^{n}\left(\frac{\partial f_i}{\partial u_i}\left(\frac{\partial u_i}{\partial \tau}\right)^2\right) < 0 \quad (9)$$

It is important to note that the HNN error function is a monotonically decreasing function, i.e., the neuron states will converge to the nearest stable point, where $\frac{du_i}{d\tau} = 0$. Consequently, the initial conditions used when solving these equations will influence the accuracy of the solution. Therefore, having a solid understanding of the underlying chemistry and physics of the problem is crucial.

The integration of equation 8 is commonly performed by numerical methods, as for example the fourth-order Runge-Kutta method.

3. Thermal Analysis

Thermal analysis can be used in a wide range of applications in chemistry and engineering,[27] such as characterization of materials,[44,45] determination of thermal stability,[46,47] investigation of kinetic mechanisms in thermal decomposition processes,[48] pyrolysis[49] and determination of adsorption parameters.[50]

The temperature control of the experiments can be set as isothermal, $T = const$, or non-isothermal, $T = T(t)$. Usually, the non-isothermal condition is programmed by linear heating rate, $\beta = dT/dt = const$, called dynamic condition.[16] This review will focus on thermogravimetry (TG) using isothermal and non-isothermal experiments and differential scanning calorimetry (DSC) with dynamic data.

For kinetic study, experimental data, i.e., mass loss for TG and heat flux for DSC, is converted to degree of conversion, α, which is adopted in the general kinetic equation[15]

$$\frac{d\alpha}{dt} = k(T)\,f(\alpha) \quad (10)$$

in which

$$\alpha = \frac{m_0 - m_i}{m_0 - m_f}(\mathrm{TG}) \text{ or } \alpha = \frac{\mathrm{H}_i}{\Delta \mathrm{H}}(\mathrm{DSC}) \quad (11)$$

with m_0, the initial mass, $m(t)$, mass along the process, m_f, final mass for TG. And H_i as the partial area calculated at time i and ΔH as the total area of the DSC signal. The rate constant, k(T), is usually given by the Arrhenius equation,[29]

$$k(T) = A\exp\left(-\frac{Ea}{RT}\right) \quad (12)$$

with A being the pre-exponential factor (usually in s^{-1} or min^{-1}), E_a, the activation energy (usually in kJ mol^{-1}) and R, the universal gas constant (8.314 J K^{-1} mol^{-1}).

Rearranging equation 10 for dynamic condition, $\beta = dT/dt = const$, the general kinetic equation is represented as,

$$\beta\frac{d\alpha}{dT} = k(T)\,f(\alpha) \quad (13)$$

In a conventional approach, equation 13 can be fitted using individual kinetic models, $f(\alpha)$, presented in Table 1. These models are proposed to describe the physical phenomenon and are grouped into 3 classes of curve profiles: acceleratory, deceleratory, and sigmoidal (or autocatalytic).[15]

By solving equation 13, it is possible to obtain the kinetic triplet, i.e., activation energy, E_a, pre-exponential factor, A, and the kinetic model, $f(\alpha)$, or its integrated function, $g(\alpha)$.[51] The A and E_a parameters can be determined by isoconversional methods,[28,52-54] generally called model-free methods, since no kinetic model is assumed in this approach.[55]

According to the International Confederation for Thermal Analysis and Calorimetry (ICTAC), the kinetics study by isoconversional methods must be performed using at least three TG curves with different β values.[15] Depending on the sample, the thermal event can occur through multiple steps with competing or consecutive reactions.[56] Thus, it is necessary to use robust methodologies to clarify the physical phenomena.[2] This work will highlight the use of neural networks to solve this chemical problem.

4. Artificial Neural Network Applied to Isothermal TG Data

Sebastião, Braga, and Yoshida[19] developed a multilayer neural network model using published experimental data for rhodium acetate. For the input, it was adopted isothermal TG curves at the temperatures 198, 202.5, 203.5, 205.0, 207.0, 209.0 and 210.5 °C.

Table 1. Kinetic models established in the literature for thermal decomposition processes[15,16]

Model	Symbol	$f(\alpha)$	$g(\alpha)$
Reaction order			
Zero-order	F1/R1	1	α
First-order	F1	$(1-\alpha)^1$	$-\ln(1-\alpha)$
Second-order	F2	$(1-\alpha)^2$	$[1/(1-\alpha)]-1$
Third-order	F3	$(1-\alpha)^3$	$(1/2)[(1-R)^{-2}-1]$
Nucleation			
Avrami-Erofeev n = 2	Am2	$2(1-\alpha)(-\ln(1-\alpha)^{1/2}$	$[-\ln(1-\alpha)]^{1/2}$
Avrami-Erofeev n = 2	Am3	$3(1-\alpha)(-\ln(1-\alpha)^{2/3}$	$[-\ln(1-\alpha)]^{1/3}$
Avrami-Erofeev n = 4	Am4	$4(1-\alpha)(-\ln(1-\alpha)^{3/4}$	$[-\ln(1-\alpha)]^{1/4}$
Geometric contraction			
Contracting cylinder	R2	$2(1-\alpha)^{1/2}$	$1-(1-\alpha)^{1/2}$
Contracting sphere	R3	$3(1-\alpha)^{2/3}$	$1-(1-\alpha)^{1/3}$
Diffusion			
1D	D1	$1/(2\alpha)$	α^2
2D	D2	$-[1/\ln(1-\alpha)]$	$((1-\alpha)\ln(1-\alpha))+\alpha$
3D: Jander model	D3	$[3(1-\alpha)^{2/3}]/[1-(1-\alpha)^{1/3}]$	$(1-(1-\alpha)^{1/3})^2$
3D: Ginstling-Brounshtein model	D4	$3/[2((1-\alpha)^{-1/3}-1)]$	$1-(2/3)\alpha-(1-\alpha)^{2/3}$
Power law			
Power Law n = 2	P2	$2\alpha^{1/2}$	$\alpha^{1/2}$
Power Law n = 3	P3	$2\alpha^{2/3}$	$\alpha^{1/3}$
Power Law n = 4	P4	$4\alpha^{3/4}$	$\alpha^{1/4}$

In this work,[19] it is proposed to replace the usual activation functions used in the MLP models by the kinetic model. Thus, equation 10 can be rewritten for constant temperature, $K(T) = k(s^{-1})$ as

$$\int_0^{\alpha} \frac{d\alpha}{f(\alpha)} = g(\alpha) = kt + k_0 \tag{14}$$

In their paper, the first order equation was used, thus the activation function for the hidden layer was defined as

$$\alpha(t) = 1 - e^{kt+k_0} \tag{15}$$

The first model studied in their paper[19] was using a single neuron in the hidden layer. By using data at different temperatures and plotting the predicted k values against $1/T$ it was possible to calculate the pre-exponential factor and activation energy, assuming the rate constant as the Arrhenius factor.

By increasing the number of neurons, prediction improvement was observed up to 8 neurons and after that it was observed overfitting of data. In this first work, the same activation function was applied to all neurons and the authors mentioned the loss of physical meaning if more than one neuron is considered.

Since the model is predicting the rate constant, k, rather than the pre-exponential factor, A, and activation energy, E_a, the results are less likely to be affected by multiple parameter combinations that yield equivalent mass loss. This issue will be discussed in the next section.

Based on the model discussed above, Sebastião *et al.*[36] developed a MLP network with a different algorithm to optimize the error function. The first step of the algorithm consists of fixing the interconnection weights between input and intermediate layers. For this, the experimental data is fitted by the individual kinect models, from Table 1, $g(\alpha) = kt + k_0$. The rate constant and linear coefficients will be represented by the interconnection weights and bias, respectively. The additive function, representing the state of neurons in the hidden layer, is determined as

$$\mathbf{W}_1\mathbf{X} = \begin{pmatrix} w_{21}t + w_{20} \\ w_{31}t + w_{30} \\ w_{41}t + w_{40} \\ w_{51}t + w_{50} \end{pmatrix} \tag{16}$$

with $\mathbf{X}$ as the state of neuron in the input layer, receiving the time data and the values w_{i1} and w_{i0} correspond to the values of the rate constants k and k_0 for the kinetic models,

respectively. Therefore, there is a physical meaning for the interconnection weights and biases in this model.

In the second stage of the algorithm, a non-linear transformation is performed, in which each neuron in the intermediate layer is activated. However, in this algorithm, each neuron in this layer is activated by a different activation function, *f*, corresponding to the kinetic models presented in Table 1, as

$$\mathbf{B} = f(\boldsymbol{W}_1\boldsymbol{X}) \tag{17}$$

The mathematical functions used in this work to activate the neurons were the Prout-Tompkins and the Avrami-Erofeev models, with m = 2, 3 and 4, described in Table 1. Thus, the neurons in the intermediate layer of the proposed MLP network are represented in equation 16. It is important to note that this network structure is not rigid, allowing the addition of other kinetic models, as well as the exclusion of some models depending on the system under study.

$$f\left(\mathbf{W}_1\boldsymbol{X}\right) = \begin{pmatrix} Au\left(\mathrm{w}_{21}\mathrm{t} + \mathrm{w}_{20}\right) \\ Am_2\left(\mathrm{w}_{31}\mathrm{t} + \mathrm{w}_{30}\right) \\ Am_3\left(\mathrm{w}_{41}\mathrm{t} + \mathrm{w}_{40}\right) \\ Am_4\left(\mathrm{w}_{51}\mathrm{t} + \mathrm{w}_{50}\right) \end{pmatrix} \tag{18}$$

Each neuron in the intermediate layer is pondered by the weights in the output layer $\mathbf{W_2}$. The vector $\mathbf{W_2}$ can be exemplified as

$$\mathbf{W}_2 = \begin{pmatrix} w_{62} & w_{63} & w_{64} & w_{65} \end{pmatrix}_{1X4} \tag{19}$$

In the third stage, a linear transformation is performed to determine the state of the neuron in the output layer, **Y**, consisting in a linear combination of the neurons in the hidden layer, as

$$\mathbf{Y} = \mathbf{W}_2 f\left(\mathbf{W}_1\boldsymbol{X}\right) \tag{20}$$

The error of the MLP network is defined considering the difference between experimental data and calculated property by the network as

$$\mathrm{E} = \left\|\mathbf{Y}_{cal} - \mathbf{Y}_{exp}\right\|_2^2, \text{ being } \mathbf{Y}_{cal} = \mathbf{W}_2 f(\mathbf{XV}) \tag{21}$$

with this network structure it is possible to calculate the contribution of the various kinetic models used in the process through a Levenberg-Marquart optimization,

$$\mathbf{W}_2 = \left(\mathbf{B}^{\mathbf{T}}\mathbf{B}\right)^{-1}\mathbf{B}^{\mathbf{T}}\mathbf{Y}_{cal} \tag{22}$$

It is interesting to note that this algorithm stands out from the others because it is a robust mathematical method in the study of thermal decomposition, since it uses the rate constants as fixed weights of the network, enabling the convergence of the network to a coherent physical result, in addition to considering the contribution of several models in the process. Therefore, this algorithm becomes useful for studying thermal decomposition of different systems. By comparing the error of the individual model adjustments with the MLP model, it was observed an increase in the adjustment by a factor of 2-7, which proves that the best description of the process is given by the combination of models rather than an individual mechanism.

This same algorithm was adopted in the MLP network developed by Ferreira *et al.*[57] on the study of anti-human immunodeficiency virus (HIV) drugs: efavirenz (EFV) and lamivudine (3TC), where it was added the kinetic models of acceleration (E potential law), deceleration (geometric model, linear, area and volume contraction), diffusion (one, two, three dimensions and Ginstling-Brounshtein). For this MLP, 13 neurons were used in the intermediate layer and the state of the neurons in the intermediate layer can be determined by

$$f\left(w_i x\right) = \begin{pmatrix} f_{D_1}\left(W_{2_1}t + W_{2_0}\right) \\ f_{D_2}\left(W_{3_1}t + W_{3_0}\right) \\ f_{D_3}\left(W_{4_1}t + W_{4_0}\right) \\ f_{D_4}\left(W_{5_1}t + W_{5_0}\right) \\ f_{R_1}\left(W_{6_1}t + W_{6_0}\right) \\ f_{R_2}\left(W_{7_1}t + W_{7_0}\right) \\ f_{R_3}\left(W_{8_1}t + W_{8_0}\right) \\ f_{A_2}\left(W_{9_1}t + W_{9_0}\right) \\ f_{A_3}\left(W_{10_1}t + W_{10_0}\right) \\ f_{A_4}\left(W_{11_1}t + W_{11_0}\right) \\ f_{F_1}\left(W_{12_1}t + W_{12_0}\right) \end{pmatrix} \tag{23}$$

In the EFV network, for the input, it was used experimental TG curves at temperatures 172, 173, 174, 175 and 178 °C. For the output, the kinetic parameters of activation energy, E_a = 75.230 kJ mol^{-1} and pre-exponential factor, A = 3.2190 × 10^{16} s^{-1} were also determined by applying the Arrhenius analysis. The residual error of

adjustment by the network is on average 10^3 times lower compared to the best individual kinetic model: R1. In the 3TC study, for the input, it was used experimental TG curves at the temperatures 215, 220, 223, 225, 228, 230, 237, 240 and 250 °C. For the output, the kinetic parameters of activation energy, E_a = 103.25 kJ mol^{-1} and pre-exponential factor, A = 2.5587 × 10^{-3} s^{-1}. The residual error of adjustment over the network is on average 10^2 times lower compared to the best individual kinetic model: D4.

This algorithm was also used in Ferreira *at al*.[58] with thalidomide, in which the MLP network also had 13 kinetic models in the intermediate layer. The activation energy determined by the network was compared with the activation energies determined along the process by the isoconversional methods of Friedman,[52] Vyazovkin,[54] and Kissinger-Akahira Sunose[53] with non-isothermal experimental data. The isothermal TG curves at temperatures 293, 304, 314, and 324 °C were used, and the best individual kinetic model was the R2 model, with the error of, on average, 10^5 times higher compared to the residual error of adjustment over the network. The MLP dealing with isothermal data presented the activation energy for the entire process of 98.94 kJ mol^{-1}. For the non-isothermal analysis, it was used four heating rates: 2.5, 5.0, 10 and 20 °C min^{-1} in N_2 atmosphere and the activation energy were calculated as 98-108 kJ mol^{-1} along the process.

Ferreira *et al*.[59] used this MLP in the study of automotive polyurethane, in which two neural networks were proposed, one being composed of eleven kinetic models and the other composed of only the Rn and Dn models, which were chosen because these models offered smaller residual error. The activation energies determined by the network were compared with those determined by the Friedman,[52] Vyazovkin,[54] and Kissinger-Akahira Sunose[53] methods with non-isothermal experimental data. The isothermal data were performed at temperatures 380, 390, 400 and 410 °C and the MLP results indicated a two-step process, in which the first occurs from the beginning until 40% of conversion, described by contraction of active nuclei, model R3, followed by diffusion, D4. The MLP adjustment of the experimental data was, on average, 10^2 to 10^3 times more accurate than the best individual kinetic model. It presented the activation energy for the entire process of 184.2 kJ mol^{-1} considering the diffusion model D4 and 198.7 kJ mol^{-1} for the contraction model. For the non-isothermal analysis, it was used four heating rates 2.5, 5.0, 10 and 20 °C min^{-1} in N_2 atmosphere, and the activation energy was calculated as 188.9 and 186.8 kJ mol^{-1}, respectively determined by the isoconversional methods Flynn-Wall-Ozawa[53] (FWO) and Kissinger-Akahira Sunose (KAS)[53] at 40% of conversion and 193.7 and 191.6 kJ mol^{-1} at extension of 60-70% of conversion.

It is important to note the necessity of adjusting each individual kinetic model before defining it as the activation function for the ANN. If these parameters are to be freely optimized while comparing the sum of different models against experimental data, different combinations of rate constants for different models could lead to the error functions with the same order of magnitude, making it difficult to evaluate which parameters have the most appropriate physical meaning.

5. Artificial Neural Network Applied to Non-Isothermal TG Data

Conesa, Caballero and Reyes-Labarta[20] developed a multilayer neural network model to predict the kinetic constants using pre-exponential factor, k_{ref}, activation energy, E_a, and reaction order, n. The model receives input simulated patterns of TG curves with 20 experimental data in each curve, calculated at different heating rates. The ANN has one hidden layer and one output layer, consisting of three neurons, one for each parameter. To generate the training data, the following reaction for pyrolysis was considered

$$F \xrightarrow{pyrolisis} (1-s)G + sS \tag{24}$$

with F being the fraction of the sample to decompose, G the gas produced during the process, S the solid residue and s the solid fraction remaining after the process has been completed. Then, the reaction speed can be determined using the global weight loss, w, as

$$\frac{dw}{dt} = -k_{ref} F^n (1-s) e^{-\frac{E_a}{R}\left(\frac{1}{T}-\frac{1}{T_{ref}}\right)} \tag{25}$$

with R the ideal gas constant. The authors mentioned that the following values were used to generate the data set: 1, 5 and 8 min^{-1} for k_{ref}; 10000, 15000, 20000, 30000 and 35000 K for $\frac{E_a}{R}$; 0.5, 1, 2 and 2.5 for n. This produced a set of 60 curves. However, it is not clear the values they have used for the heating rates, F, T_{ref} and s while generating the training data.

For training the model with theoretically calculated values, the authors splitted the data into different subsets, one for training (used for updating the model weights) and validation (to evaluate the model performance on unseen data during the training procedure). They have used two different approaches to avoid overfitting: (*i*) regularization,

which applies a penalty to the error function if large weights are used during the training; (*ii*) early stopping, which stops the training procedure if the validation data error starts increasing. They determined that 10 neurons were used in the input layer for the regularization approach and 4 for the early stopping, both using sigmoid as the activation function. In the trained model, it is not clear which one of the two models mentioned, was able to retrieve the constants in agreement with the values acquired using other methods while using experimental data for cellulose, polyethylene and lignin. It is emphasized by the authors that the model is specific for systems that can be modeled by single n^{th}-order reactions.

Caution must be taken while simulating data with pre-selected E_a and k^{ref} values for training an ANN model. In equations such as (equation 25), the factor $K(T) = k_{ref} e^{\frac{E_a}{R}\left(\frac{1}{T}-\frac{1}{T_{ref}}\right)}$ may have the same value for different combinations of k_{ref} and E_a, this is known as kinetic compensation effect.[60]

Therefore, the parameter selected may bias the model predictions when applying it to experimental data. Although this procedure does not allow for a straightforward evaluation of the predicted individual parameters, the final value for $K(T)$ is likely to be the same for all model predictions. Different combinations of parameters that lead to the same similar exponential factor have been also observed for another problem.[61]

Therefore, if the objective is to determine the rate constant for the reaction, ANN models would offer more reliability, as discussed in the section 4. However, if the objective is obtaining E_a and k_{ref}, that should be done using isoconversional or Arrhenius analysis and using the neural network for making other predictions. If ANN models are used to predict these parameters, they should be trained on a dataset containing multiple experimental curves where the values for these parameters were obtained through the analysis mentioned previously. This approach ensures that the model learns from real data, enhancing the reliability of its predictions.

ANN can be used to model different properties of a material and provide insights about its formulation. Burgaz *et al.*[17] investigate the use of artificial neural networks (ANNs) to predict the thermal stability, crystallinity, and thermomechanical properties of different proportions of poly(ethylene oxide) (PEO) and clay nanocomposite compositions. The proportions obtained and simulated were PEO – n% clay, with n = 0, 5, 10, 15, 20, and 100. The authors trained different ANNs for each type of experimental data: thermogravimetric analysis (TGA), DSC, and dynamic mechanical analysis (DMA). All input data were normalized using the normalization equation,

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}}(\text{high limit} - \text{low limit}) + \text{low limit} \quad (26)$$

with x_{min} and x_{max} the minimum and maximum values of experimental data. High and low limits are defined as 1 and 0, respectively. All neural networks used the feed-forward back propagation algorithm. The data was splitted into training (75%), validation (15%), and testing (15%) sets. For DMA neural networks, one hidden layer was enough to predict the storage modulus (E') and tan δ (the ratio of the loss modulus (E'') to the storage modulus (E')) using 193 training data for each case. For TGA data, a neural network was necessary to have two hidden layers to predict weight loss using 406 training data. For DSC data, one hidden layer was used to predict the heat flow using 565 training data. The results demonstrate the potential of ANNs for predicting different properties of polymer/clay nanocomposites. The training process revealed that increasing the amount of data enhances the network's modeling capacity. In DMA analysis, the relationship between the storage modulus and temperature is much less nonlinear compared to relations between the input and output parameters in both DSC and TGA. Thus, the modeling capability of ANN is the highest for the DMA analysis. Therefore, the authors increased the number of hidden layers to improve the accuracy for TGA and DSC data. An option broadly discussed to improve the accuracy in modeling thermal data is the use of more than one heating rate, which could be another possibility for the performance enhancement in this work, as can be seen in ICTAC.[25,27] The neural networks developed in this work do not provide information about the kinetic parameters and mechanisms for the experimental data.

Muravyev and Pivkina[13] proposes the application of ANNs for predicting kinetic parameters from simulated thermoanalytical data. The input data consisted of conversion degree *versus* temperature profiles ($\alpha(T)$) generated for single and multiple heating rates, as well as constant-rate thermal analysis (CRTA) runs. The output included the kinetic triplet (activation energy E_a, pre-exponential factor $\log A$, and reaction model $f(\alpha)$) with ANNs trained on a broad range of parameters ($E_a = 50 - 300$ kJ mol^{-1}, $\log A = 0.01 - 40$ s^{-1}) and 10 reaction models.

Key findings demonstrate that ANNs, particularly multilayer perceptrons, achieve high accuracy (performance > 99% for some models) in predicting kinetic parameters, outperforming traditional single-heating-rate methods. Advantages of the ANN approach include rapid analysis, tolerance to noisy data, and adaptability to complex kinetic scenarios. However, limitations were noted for single-heating-rate inputs (low predictive power outside

training ranges) and simultaneous prediction of multiple parameters (reduced accuracy). The work underscores ANNs as a promising alternative to conventional kinetic methods. A critical point highlighted was that the simulated data considered was single-step reactions, considering that greater computational power is required to work with multi-step reactions.

In reference,[13] a broad range of parameters selection for generating the data set is a good practice, since it increases the amount of data for training the model. The authors mentioned the kinetic compensation effect while trying to determine the kinetic triplet in the usual methods and that, hypothetically, using ANN would result in lower parameter interdependency. However, the kinetic compensation effect may still be a problem while generating the simulated data. For all combinations of pre-exponential factors and activation functions considered, it would be good to have an analysis of the set of parameters that would produce the same reaction rates. Thus, although the model had an excellent performance, the evaluation of these parameter predictions remains a challenge.

Rojek, Suchacz, and Wesolowski[62] proposed performing incompatibility detection between active pharmaceutical ingredients (API) and excipients based on thermogravimetric data using a self-organization map (SOM). In this work, binary mixtures of caffeine and selected excipients were used as a study model.

The SOM network can be applied for clustering, feature extraction and anomaly detection. It will map higher dimensional data into a lower dimensional grid. In this type of model, the input data is passed to a grid of interconnected neuros and one of these will be promoted to an activated state. Thus, similar data will activate neurons close to each other.

The input data for this model is a discrete set of temperatures corresponding to a percentage of mass loss, from 5% (T_5) up to 75% (T_{75}) in increments of 5%. To appropriately map the data, different models consisting of 2D grids were tested, 3 × 3, 4 × 4, 5 × 5 and 6 × 6, and the 5 × 5 grid was found to be the optimal for this problem.

After training the model, the weight values of the neuron connections were used to perform a principal component analysis (PCA) and cluster analysis (CA) to identify the temperature of mass loss which would potentially indicate incompatibility. The predictions of the model were compared against DSC, Fourier transform infrared spectroscopy (FTIR) and powder X-ray diffraction (PXRD). It was found that DSC and FTIR agreed with the predictions and that the range of temperatures potentially most important for indicating incompatibility are those corresponding to mass losses from 20 to 55%.

Therefore, this result indicated that ANNs combined with thermogravimetry can be used as a diagnostic tool to determine incompatibility occurrence between (API) and excipients.

Özge Çepelioğullar *et al.*[63] applied ANNs to predict refuse-derived fuel (RDF) pyrolytic behaviors. The model proposed in their study consists of one input layer that receives temperature and heating rate data, two hidden layers and an output layer that predicts the weight loss at that temperature.

The data was generated using 7 different heating rates: 5, 10, 15, 20, 30, 40 and 50 °C min^{-1}. For each parameter it was recorded approximately 4000 datapoints. From this data set, the optimal number of 1000 points to be selected at random was acquired by trial and error. From the selected data, 75% was used for training, 15% for testing and 15% for validation steps. The training process was carried out for different combinations parameters: number of neurons in the hidden layer varying from 3 to 7, giving a total of 25 possible combinations; activation functions used in the hidden layer were tan-sigmoidal or log-sigmoidal, giving a total of 4 different combinations; each model was trained 5, 10 and 25 times, which gives 3 different combinations. Thus, in total, there were 100 different models trained 3 times each. The best performance was observed for the model trained 25 times, containing 7 and 6 neurons in the first and second hidden layers, with tan-sigmoidal or log-sigmoidal activation functions, respectively.

In their work,[63] the model was able to predict the data with high accuracy for all training, test and validation sets. The model was used to predict a new TG curve generated with a heating rate of 25 °C min^{-1} and the predictions were very accurate while comparing against experimental data. This work makes evident the importance of the model design step. It is important to note at this point that a good and generalizable model should be able to predict data within the limits of the data used for the training step, in this case for heating rates of 5 and 50 °C min^{-1} and minimum/max temperatures. The authors highlight the effectiveness of ANNs in capturing the complex thermal decomposition characteristics of RDF, suggesting their potential as powerful tools for reducing the number of TG experiments needed.

Monticeli *et al.*[64] uses an artificial neural network to predict thermal degradation curves for four different fibers: ramie, curaua, kenaf and jute. While describing the ANN approach, it is mentioned the Conesa *et al.*[20] work and described a process for data acquisition assuming the same n^{th} order kinetic reaction given by equation 25. However, there is no further mention to this theoretical dataset in the text.

The model proposed in their work[64] receives as inputs the time, temperature and heating rate. It has 12 hidden layers and outputs the prediction for the mass loss. The data set was acquired experimentally with temperature ranging from 25 to 800 °C and using heating rates of 5, 10, 20 and 40 °C min^{-1}.

The authors[64] started the ANN model training by first considering only one heating rate, 10 °C min^{-1}, and they separated the data into training and test sets and the method performance was evaluated based on the determination coefficient of the test set. The number of data to be used in the training step was determined by evaluating the model performance for all four fibers for different numbers of training data. However, it is not clear the total number of data nor what percentage the number of training data represents. They have observed that as the number of training data increases, the performance for all fibers analyzed converged to the same value for number of training data of 60. Without knowing what percentage this number represents, it is hard to evaluate this selection as appropriate or not. The regression plots presented for comparing predicted and target values shows an excellent agreement.

After describing the training process, it is said that the same procedure was performed for the other heating rates, to achieve an ANN model capable of prediction data at heating rates not accessed experimentally. However, it is not clear what procedure the authors have used: did they use transfer learning for each one of the other heating rates? Or did they retrain the model using all data at the same time using the hyperparameters determined in the first step? The performance for unseen heating rates is demonstrated by showing a plot for the three-dimensional TG surface having 5 and 40 °C min^{-1} as lower and upper limits.

Still regarding the training step, in the conclusion section the authors said only one heating rate was used to train the model. This seems to contradict the description in the results section.

Araujo *et al.*[65] investigated the application of MLP neural networks to the biopolymer chitosan's TG dynamic data. The authors aim to demonstrate the effectiveness of MLP neural networks in modeling the kinetics of solid thermal decomposition. For the chitosan experimental data, the MLP shows the lowest residual error to fit experimental data, which were performed at N_2 atmosphere under four heating rates: 2.5, 5.0, 7.5, and 10 °C min^{-1}. The activation energy was calculated as 98.1 to 183.3 kJ mol^{-1} along the conversion degree. The findings show that MLP neural networks provide a robust and accurate tool for understanding the complex decomposition behaviors of biopolymers, enhancing the predictive capability for thermal analysis in various industrial applications. This methodology has been applied to different materials' thermal decomposition, for example, graphene oxide,[22] nitroxides,[66] and Eu^{3+}-betadiketonate coordination compounds.[67]

The model architecture is based on the isothermal neural network proposed by Sebastião *et al.*,[36] and follows the same development discussed in the section 4. The innovation consists of adding neurons in the input layer, representing the temperature for each conversion degree in different heating rates, β. Also, another difference is the interconnection weights between the intermediate and input layers, which are determined assuming,

$$W_1\left(T_{\alpha i}\right)=\left(\frac{AR}{\beta}\frac{}{E_\alpha}T^2 exp\left(-\frac{E_\alpha}{RT_{\alpha i}}\right)\right) \quad (27)$$

with the activation energy and pre-exponential factor determined using the isoconversional method.[28,54] The kinetic models presented in Table 1 also activate the hidden layer, thereby ensuring not only the optimization of the neural network but also preserving the physical significance of the weights.

De Freitas-Marques *et al.*[68] also applied the MLP neural network proposed above for the kinetic studies of the drugs atazanavir sulfate (ATA) and ritonavir (RIT) thermal analyses. The thermal behavior of ATA indicates melting at 190.7 °C (ΔH_m = 71.5 J g^{-1}) with decomposition (E_a = 101.8 kJ mol^{-1}). For RIT indicates melting at 119.2 °C (ΔH_m = 66.5 J g^{-1}) with decomposition (E_a = 100 kJ mol^{-1}; α = 10%). The kinetic model that contributes the most is the Avrami-Erofeev model with m = 2 and describes the thermal decomposition of ATA and RIT. The MLP model for atazanavir was trained using the heating rate curves at 14, 18, and 20 °C min^{-1}. It was shown that the kinetic parameters, E_a, A, and the kinetic contribution model retrieved for the MLP neural network were able to recover an experimental curve at 40 °C min^{-1}. It demonstrates the capability of the proposed neural network to extrapolate experimental data.

Araujo *et al.*[56,69] explored using the HNN to model kinetic behaviors and activation energy distribution in complex systems. For this approach, it is assumed that several reactions may occur in the same event and the activation energy, for the whole process, is described by a probability density function, $f(E_a)$, instead of an average value, where each reaction has its probability in the process, $P(E_a)$. For this, the experimental data can be represented as,

$$\left(1-\alpha\right)=\sum P\left(\mathrm{E_a}\right)e^{-\int\frac{k}{\beta}dT} \quad (28)$$

The probability function is determined by $P(E_a) = \frac{dV}{V^*} = \int_{E_a}^{E_a+dE_a} f(E_a)dE_a$. The integral can be further developed as $\int f(E_a)dE_a = \sum f(E_{ai})\Delta(E_{ai}) = \sum \frac{P(E_{ai})}{\Delta(E_{ai})}\Delta(E_{ai})$, in which the $\Delta(E_{ai})$ is a convenient range of E. Considering a continuous distribution of activation energy and also the probability density function, $f(E_a) = P(E_a)/\Delta(E_a)$, the equation 28 can be rewritten as equation 29 and can be interpreted as the first-order integral Fredholm equation,

$$(1-\alpha) = \sum_{i=1}^{n} w_i \int_0^{\infty} K(T, E_a) f_i(E_a) dE_a \quad (29)$$

with $K(T, E)$ the kernel and $f(E_a)$ the activation energy distribution function. This equation represents a class of problems known as ill-posed inverse problem, and the HNN approach, tested with both simulated and experimental data, demonstrates better efficiency compared to the traditional methods, offering a promising tool for thermal analysis in industrial applications. Araujo *et al.*[56,69] used the HNN to estimate kinetic parameters in the distributed activation energy models (DAEM). This method can estimate the shape of $f(E)$ without *a priori* assumption of its functional form, and a constant pre-exponential factor, A. For training the model, simulated data with random noise was performed. For the kinetic model determination, the HNN algorithm was adapted to obtain the parameters of the Cai-Liu model, $f(\alpha) = \alpha^m(1 - q\alpha)^n$. In addition, their model was robust to noise in the experimental data.

Wakimoto *et al.*[70] proposed an improvement for the method proposed by Araújo *et al.*,[56] the new methodology estimates the kinetic parameters in the DAEM without *a priori* assumption of both $f(E_a)$ and the pre-exponential factor A. The model was tested by performing a kinetic analysis on reaction data of the parallel reaction system generated by numerical simulations under seven heating conditions in which N is the reactions proceed in parallel, and the data can be written as,

$$(1-\alpha) = \sum_i^N \frac{V_i^*}{V_i} \exp\left[-\exp\left(-\ln\frac{\beta}{T^2} - \frac{E_{ai}}{RT} + \ln\frac{A_i R}{E_{ai}}\right)\right] \quad (30)$$

where V_i^*/V_i denotes the contribution of the ith reaction. The estimation of conversion α, can be regarded as the forward propagation of the three-layer neural network and the kinetic parameters A_i, E_{ai}, and V_i^*/V_i can be obtained as the optimized weights and biases of this neural network.

For that, $\ln\frac{\beta}{T^2}$ and 1/T are the input to the neural network and are normalized by Z-score normalization[71] to prevent a negative influence of the input variables on the optimization using the gradient descent method and the loss function to be minimized is defined by,

$$(objective) = \frac{1}{M}\sum_{j=1}^{M}\left(X_j - X_{exp,j}\right)^2 + \lambda\left|\left(\sum_{i=1}^{N}\frac{V_i^*}{V_i} - 1\right)\right| \quad (31)$$

where X_{exp} is the experimental value of conversion and M, the number of data points. The first term in equation 31 is the mean squared error (MSE), and the second term is used to constrain the sum of V_i^*/V^* to unity, with $\lambda = 0.1$ being a coefficient that balances the first and second terms.

The limitation of the proposed method is the determination of the hyperparameters in the neural network. The proposed method has several hyper-parameters, and optimizing all of them requires too much computational cost. In this study, the authors changed the lower limit of V_i^*/V^* , and fixed the other hyperparameters. The number of hidden layer nodes in the neural network (i.e., the number of reactions proceeding in parallel) was set to N = 64. The possible range of E_a was limited to 0.1-10000 kJ mol^{-1}. The initial activation energy of each reaction was set randomly in E_a = 0.1-400 kJ mol^{-1}. For the gradient descent method, the number of optimization steps was set to 10^5.

6. Artificial Neural Network Applied to DSC Data

For the DSC analysis, Sbirrazzuoli and Brunel[23] showed the mechanism for homogeneous kinetic processes as $f(\alpha_i) = (1 - \alpha_i)^n$ is appropriate for many systems.[31,72] Assuming the equation 10, whit $\alpha_i = H_i/\Delta H$ being determined from H_i, as the partial area calculated at time i, and ΔH as the total area of the DSC signal, one has

$$\left(\frac{dH}{dt}\right)_i = \Delta H \cdot A e^{-\frac{E_a}{RT_i}} \cdot f(\alpha_i) \quad (32)$$

in which the DSC experimental data of heat flow is $P_i = (dH_i/dt)$. By taking the logarithm in equation 32, we can establish the general equation,[73,74]

$$ln\left[\left(\frac{dH}{dt}\right)_i\left(\frac{1}{\Delta H}\right)\right] - n\ln(1-\alpha_i) = lnA - \frac{E_a}{RT_i} \quad (33)$$

Sbirrazzuoli and Brunel[23] were the first to propose using artificial neural networks for filtering and deconvoluting calorimetric signals.[31] Their study utilized synthetic thermoanalytical curves of heat flow relative to temperature, with the objective function being the difference between the synthetic data, $Y_{r,i,exp}$ and the recovered curve $Y_{r,i,cal}$ as

$$E = \sum_{r=1}^{M} \frac{1}{N} \sum_{i=1}^{N} \left(Y_{r,i,exp} - Y_{r,i,cal} \right)^2 \quad (34)$$

with N as the number of recorded points and M as the number of patterns. The aim of the study was to evaluate the ability of neural networks to determine thermodynamic and kinetic parameters from DSC curves.

The training data was simulated by solving equation 33 assuming $f(\alpha_i) = (1 - \alpha_i)^n$, with $n = 2$, activation energy ranging from 74 to 80 kJ mol^{-1} and ln(A) ranging from 18 to 20. These values were based on previous experimental results for the reaction of interest. In reference 23, the authors mentioned that after generating the data, the kinetic parameters and enthalpy are computed using the single peak method. In reference 31, it seems the values used as target were those used for simulating the data, without carrying out the parameter recovery step.

The optimal topology for their ANN model was determined as receiving 320 input values, representing the DSC curve, having three hidden layers with 15, 20 and 20 neurons and outputting the results for ΔH, E_a and ln(A). For the training step, the authors used the resilient-propagation algorithm[23] to optimize the model weights. The total number of curves to be used for training the model was evaluated by carrying out the process for different steps for E_a and ln(A). Also, noise was added to the training input data to assess the model capability to learn and sensibility to noisy data as input.

The authors in reference 31 found that the model was able to recover the parameters used for simulating the curves with good accuracy with both clean and noisy data. Demonstrating that ANN models are potential useful tools for thermal analysis.

It is interesting to note that for their work, the range of values used for generating the data is based on previously determined experimental results. That potentially makes the model prediction more reliable. Also, the model was not applied to experimental data for making predictions. If that was the case, it is always important to note that using simulated data for training a model makes it difficult to assess if the model is learning physically meaningful trends or artifacts due to parameter correlations, as discussed previously about the kinetic compensation effect. The authors also mentioned about the possibility of finding different combinations of parameters corresponding to low errors while using single-peak methods.[75]

Considering another strategy, equation 2 was treated by Hopfield neural network-based algorithm by Sebastião and co-workers[76] to determine the kinetic parameters m and n as the reaction order, E_a the activation energy, and A the pre-exponential factor. A general kinetic model $f(\alpha) = \alpha^m(1 - q\alpha)^n$ was considered, in which the mechanism depending on the m, n and q parameters, describing the physics and chemistry of the process.[30,73]

The Hopfield neural network-based algorithm used the equation 26 as the multi objective error function and to investigate the performance of the algorithm to fit DSC data and compute the kinetic parameters. It was used 1000 synthetic data of $dH/dt(T)$, for each heating rate, and to each case it was incorporated random errors of 2.5% point-by-point in the $Y_{r,i,exp}$ synthetic curves. It was observed that the HNN method is not sensitive to noisy data.

Experimental data of an antihypertensive drug, losartan potassium, LOK, during its polymorphic conversion were investigated.[76] It was assumed

$$rmsd = \sum_{r=1}^{3} \sqrt{\frac{\sum_{i=1}^{N} \left(\left(\frac{dH}{dt} \right)_{r,i(calc)} - \left(\frac{dH}{dt} \right)_{r,i(exp)} \right)^2}{N}} \quad (35)$$

with $rmsd$ as the root mean square deviation and three heating rates of DSC data in a multi objective error function. From the results, it was stated the LOK conversion presented $A = 6.376 \times 10^{12}$ s^{-1}, $E_a = 136.5$ kJ mol^{-1}, $m = 0.6$, $n = 0.85$ and $q = 1.0$. Therefore, it is observed that the process occurs as a combined event, once the values of these parameters in the $f(\alpha)$ function do not correspond to any ideal kinetic model.

When using the HNN method, it is important to keep in mind that the method converges to the nearest stable solutions. Therefore, when defining the initial guess to the set of Hopfield different equations, it is important to have a good estimate of the values for the real process. Additionally, using simulated data is not an issue, as it does not rely on training with a dataset before being applied to experimental data.

In this work, the authors have tested different initial conditions while applying the method to experimental data and found different sets of parameters that lead to small residuals. In that case, since all parameters were physically acceptable, the set with smaller residual was selected as the final answer.

7. Conclusion

This review highlights the potential of artificial neural networks (ANNs) in analyzing and interpreting thermal analysis data, specifically TG and DSC. A table with the papers discussed is available in the Supplementary Information section. The application of ANNs, with emphasis on the MLP and HNN, has significantly enhanced the accuracy and complexity of kinetic analyses for both

isothermal and non-isothermal processes. By integrating simulated and experimental data with robust computational algorithms, these networks provide a powerful framework for predicting thermal behavior, characterizing kinetic parameters, and modeling complex reaction mechanisms. The findings demonstrate that ANN-based methodologies surpass traditional optimization algorithms like Levenberg-Marquardt in adaptability and robustness, being able to handle noise in experimental data and complex multi-step reaction processes.

However, some caution must be taken while conducting studies in this field. For models such as MLP, that rely on a training data set, using simulated data during the training step previously to the application to experimental data, for predicting kinetic parameters, may introduce bias in the model predictions, as different combinations of pre-exponential factor and activation energy can lead to the same result. For training these models, ideally it should be used an extensive experimental dataset from which the kinetic parameters were extracted from other methods, as, for example, the isoconversional analysis. As for the HNN, it is important to have a good initial guess for the problem solution to prevent the method from converging to a minimum on the error surface that corresponds to non-physical parameters. The existence of multiple minima on the error surface can also be an issue for MLP models. Therefore, it is important to train and test the model with the chosen topology using multiple different initialization for weights and biases to ensure that the topology can learn the problem effectively, regardless of its initial parameters.

Future work should explore the broader applicability of these methodologies to diverse materials and conditions. Additionally, advances in machine learning architectures and training algorithms can possibly provide further improvement of its predictive capabilities and computational efficiency of ANNs in thermal analysis. The integration of these technologies represents a paradigm shift, bridging experimental thermal analysis with computational intelligence to unlock new possibilities in materials science and engineering.

Supplementary Information

Supplementary data are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

The authors would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), process 05870/2023-8, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES) for financial support, Programa de Pós-graduação em Química da UFMG (PPGQUI-UFMG) and Programa de Pós-graduação em Inovação Tecnológica da UFMG (PPGIT-UFMG).

Author Contributions

Rita C. O. Sebastião was responsible for the conceptualization, writing original draft, review and editing, project administration; Natália R. S. Araujo for the conceptualization, writing original draft, review and editing; Felipe S. Carvalho for writing-review and editing; Bárbara D. L. Ferreira for the conceptualization and writing original draft; João Pedro Braga for writing-review and editing.

Rita C. O. Sebastião *is a full professor at Universidade Federal de Minas Gerais (UFMG) with experience in Chemistry, working mainly on the topics: inverse problems, artificial neural networks, nuclear magnetic resonance, thermal decomposition of solids, chemical kinetics, combustion processes, and thermodynamics. Extensive experience in education, research, and extension, especially in projects aiming for the popularization of science, educational innovation, and entrepreneurship, connecting the university to basic education. Experience in developing and managing projects on socio-environmental sustainability and education. She is currently a member of the Postgraduate Program in Chemistry, Postgraduate Program in Technological Innovation and Professional Master's in Technological Innovation and Intellectual Property at UFMG.*

Natália R. S. Araujo *is a PhD candidate in Chemistry at Universidade Federal de Minas Gerais, where she also earned her Bachelor's and Master's degrees. She has experience in Physical Chemistry, focusing on the development of MLP and Hopfield neural networks, combustion kinetics, and thermal decomposition kinetics. Her research includes theoretical chemistry and thermal analysis.*

Felipe S. Carvalho *earned his bachelor's in Chemistry from Universidade Federal de Viçosa and his MSc and PhD in Theoretical Chemistry from Universidade Federal de Minas Gerais. His research interests include inverse problems, chemical kinetics, adsorption, deep learning applied to*

physical chemistry, and statistical thermodynamics of liquids. As a research scholar at California State University, Northridge, he contributed to 3D-RISM method development in AmberTools: including CUDA acceleration, ionic parameter optimization and a water placement package. Currently, he is a research scholar at the California Institute of Technology, studying molecular dynamics of polymers at water-surface interface.

Bárbara D. L. Ferreira *is Bachelor (2010) in Chemistry from the Universidade Federal de Viçosa. Master (2012) and Doctor (2016) in Chemical Sciences from the Universidade Federal de Minas Gerais. Specialist in Higher Education (2020) and Chemistry Teaching (2021) from Claretiano Centro Universitário (2020). She has postgraduate degree in Teacher Training for EaD (2020) and Educational Design (2020) from CEFOR-Training Center of Instituto Federal do Espírito Santo. She has experience in Physical Chemistry, ill-conditioned inverse problems, Hopfield artificial neural network and MLP, combustion kinetics, thermal decomposition kinetics, nuclear magnetic resonance and image reconstruction on computed tomography.*

João Pedro Braga *holds a degree in Chemistry (1979), a master's degree in Physics (1982) both from the Federal Universidade Federal de Minas Gerais (UFMG), a PhD in Chemistry from the University of Sussex (1986) and a Visiting Fellow position at Princeton University (1996). He is currently a full professor (retired) at UFMG. He has published seven books (Physical Chemistry, Quantum Chemistry, Statistical Thermodynamics, Marie Curie's visit to Brazil, Maria Sklodowska-Curie W Brazylii, Statistical Thermodynamics/second edition, Numerical Methods in Quantum Chemistry) and spent thirty-seven years as a productivity fellow at CNPq. Professor Braga has experience in the area of Physical Chemistry.*

References

1. Malik, S.; Muhammad, K.; Waheed, Y.; *Ain Shams Engineering J.* **2024**, *15*, 102886. [Crossref]
2. Muravyev, N. V.; Luciano, G.; Ornaghi, H. L.; Svoboda, R.; Vyazovkin, S.; *Molecules* **2021**, *26*, 3727. [Crossref]
3. Jordan, M. I.; Mitchell, T. M.; *Science* **2015**, *349*, 255. [Crossref]
4. Haykin, S.; *Neural Networks: A Comprehensive Foundation*, 3rd ed.; Prentice-Hall, Inc.: Upper Saddle River, NJ, 2007.
5. Durodola, J. F.; Li, N.; Ramachandra, S.; Thite, A. N.; *Int. J. Fatigue* **2017**, *99*, 55. [Crossref]
6. Hossain, M. A.; Rahman, M. M.; Prodhan, U. K.; Khan, M. F.; *arXiv:1308.3785* **2013**. [Crossref]
7. Świetlicka, I.; Kuniszyk-Jóźkowiak, W.; Świetlicki, M.; *Sensors* **2022**, *22*, 321. [Crossref]
8. Goldberg, Y.; *Neural Network Methods for Natural Language Processing*; Morgan & Claypool: Toronto, 2017. [Crossref]
9. Cho, B. C.; Liaw, S.-L.; Chang, C.-N.; Yu, R.-F.; Yang, S.-J.; Chiou, B.-R.; *Water Sci. Technol.* **2001**, *44*, 95. [Crossref]
10. Kaastra, I.; Boyd, M.; *Neurocomputing* **1996**, *10*, 215. [Crossref]
11. Brown, M. E.; Gallagher, P. K.; *Handbook of Thermal Analysis and Calorimetry: Recent Advances, Techniques and Applications*, vol. 5; Elsevier: Amsterdam, 2011.
12. Rudin, C.; *Nat. Mach. Intell.* **2019**, *1*, 206. [Crossref]
13. Muravyev, N. V.; Pivkina, A. N.; *Thermochim. Acta* **2016**, *637*, 69. [Crossref]
14. Galván, I. M.; Zaldívar, J. M.; Hernández, H.; Molga, E.; *Comput. Chem. Eng.* **1996**, *20*, 1451. [Crossref]
15. Vyazovkin, S.; Burnham, A. K.; Criado, J. M.; Pérez-Maqueda, L. A.; Popescu, C.; Sbirrazzuoli, N.; *Thermochim. Acta* **2011**, *520*, 1. [Crossref]
16. Khawam, A.; Flanagan, D. R.; *J. Phys. Chem. B* **2006**, *110*, 17315. [Crossref]
17. Burgaz, E.; Yazici, M.; Kapusuz, M.; Alisir, S. H.; Ozcan, H.; *Thermochim. Acta* **2014**, *575*, 159. [Crossref]
18. Al-Yaari, M.; Dubdub, I.; *Polymers* **2020**, *12*, 1813. [Crossref]
19. Sebastião, R. C. O.; Braga, J. P.; Yoshida, M. I.; *J. Therm. Anal. Calorim.* **2003**, *74*, 811. [Crossref]
20. Conesa, J. A.; Caballero, J. A.; Reyes-Labarta, J. A.; *J. Anal. Appl. Pyrolysis* **2004**, *71*, 343. [Crossref]
21. Vieira, S. S.; Paz, G. M.; Araujo, B. C. R.; Lago, R. M.; Sebastião, R. C. O.; *Thermochim. Acta* **2020**, *689*, 178628. [Crossref]
22. Menezes, I. R. S.; Araújo, N. R. S.; Araújo, B. C. R.; Sakai, T.; Lago, R. M.; Sebastião, R. C. O.; *Thermochim. Acta* **2023**, *721*, 179454. [Crossref]
23. Sbirrazzuoli, N.; Brunel, D.; *Neural Comput. Applic.* **1997**, *5*, 20. [Crossref]
24. Koga, N. In *Handbook of Thermal Analysis and Calorimetry*, vol. 6, 2nd ed.; Elsevier: Amsterdam, 2018, ch. 6. [Crossref]
25. Vyazovkin, S.; Burnham, A. K.; Favergeon, L.; Koga, N.; Moukhina, E.; Pérez-Maqueda, L. A.; Sbirrazzuoli, N.; *Thermochim. Acta* **2020**, *689*, 178597. [Crossref]
26. Braga, J. P.; Lemes, N. H. T.; Borges, E.; Sebastião, R. C. O.; *Quim. Nova* **2016**, *39*, 886. [Crossref]
27. Koga, N.; Vyazovkin, S.; Burnham, A. K.; Favergeon, L.; Muravyev, N. V.; Pérez-Maqueda, L. A.; Saggese, C.; Sánchez-Jiménez, P. E.; *Thermochim. Acta* **2023**, *719*, 179384. [Crossref]
28. Vyazovkin, S.; *J. Therm. Anal.* **1997**, *49*, 1493. [Crossref]

29. Vyazovkin, S.; Wight, C. A.; *Thermochim. Acta* **1999**, *340-341*, 53. [Crossref]
30. Vyazovkin, S.; Chrissafis, K.; Di Lorenzo, M. L.; Koga, N.; Pijolat, M.; Roduit, B.; Sbirrazzuoli, N.; Suñol, J. J.; *Thermochim. Acta* **2014**, *590*, 1. [Crossref]
31. Sbirrazzuoli, N.; Brunel, D.; Elégant, L.; *J. Therm. Anal.* **1997**, *49*, 1553. [Crossref]
32. Mahesh, B.; *Int. J. Sci. Res.* **2020**, *9*, 381. [Crossref]
33. Wu, Y.; Feng, J.; *Wireless Pers. Commun.* **2018**, *102*, 1645. [Crossref]
34. McCulloch, W. S.; Pitts, W.; *Bull. Math. Biophys.* **1943**, *5*, 115. [Crossref]
35. Apicella, A.; Donnarumma, F.; Isgrò, F.; Prevete, R.; *Neural Networks* **2021**, *138*, 14. [Crossref]
36. Sebastião, R. C. O.; Braga, J. P.; Yoshida, M. I.; *Thermochim. Acta* **2004**, *412*, 107. [Crossref]
37. Chauvin, Y.; Rumelhart, D. E.; *Backpropagation: Theory, Architectures, and Applications*; Lawrence Erlbaum Associates: Hillsdale, New Jersey, USA, 1995.
38. Suliman, A.; Zhang, Y.; *J. Earth Sci. Eng.* **2015**, *5*, 52. [Link] accessed in May 2025
39. Hopfield, J. J.; *Proc. Natl. Acad. Sci.* **1982**, *79*, 2554. [Crossref]
40. Hopfield, J. J.; *Proc. Natl. Acad. Sci.* **1984**, *81*, 3088. [Crossref]
41. Hopfield, J. J.; Tank, D. W.; *Biol. Cybernetics* **1985**, *52*, 141. [Crossref]
42. Lemes, N. H. T.; Borges, E.; Braga, J. P.; *J. Braz. Chem. Soc.* **2007**, *18*, 1342. [Crossref]
43. Sebastião, R. C. O.; Lemes, N. H. T.; Virtuoso, L. S.; Braga, J. P.; *Chem. Phys. Lett.* **2003**, *378*, 406. [Crossref]
44. Silva, C. R. G.; Fialho, S. L.; Barbosa, J.; Araújo, B. C. R.; Carneiro, G.; Sebastião, R. C. O.; Mussel, W. N.; Yoshida, M. I.; de Freitas-Marques, M. B.; *J. Braz. Chem. Soc.* **2021**, *32*, 638. [Crossref]
45. Peñalver, R.; Arroyo-Manzanares, N.; López-García, I.; Hernández-Córdoba, M.; *Chemosphere* **2020**, *242*, 125170. [Crossref]
46. Reda, S. Y.; *Ciência e Tecnologia de Alimentos* **2011**, *31*, 475. [Crossref]
47. Yan, Q.-L.; Zeman, S.; Elbeih, A.; *Thermochim. Acta* **2012**, *537*, 1. [Crossref]
48. L'vov, B. V; *Thermochim. Acta* **2002**, *386*, 1. [Crossref]
49. Slopiecka, K.; Bartocci, P.; Fantozzi, F.; *Appl. Energy* **2012**, *97*, 491. [Crossref]
50. Senthil Kumar, P.; Senthamarai, C.; Durgadevi, A.; *Environ. Prog. Sustainable Energy* **2014**, *33*, 28. [Crossref]
51. Schick, C.; Vyazovkin, S.; Koga, N.; *Handbook of Thermal Analysis and Calorimetry*, vol. 6, 2nd ed.; Elsevier: Amsterdam, 2018. [Crossref]
52. Friedman, H. L.; *J. Polym. Sci. B* **1969**, *7*, 41. [Crossref]
53. Mothé, C. G.; de Miranda, I. C.; *J. Therm. Anal. Calorim.* **2013**, *113*, 497. [Crossref]
54. Vyazovkin, S.; Dollimore, D.; *J. Chem. Inf. Comp. Sci.* **1996**, *36*, 42. [Crossref]
55. Arenas, C. N.; Navarro, M. V.; Martínez, J. D.; *Bioresour. Technol.* **2019**, *288*, 121485. [Crossref]
56. Araújo, N. R. S.; Duarte, Á. C. M.; Pujatti, F. J. P.; Freitas-Marques, M. B.; Sebastião, R. C. O.; *Thermochim. Acta* **2021**, *697*, 178847. [Crossref]
57. Ferreira, B. D. L.; Araujo, B. C. R.; Sebastião, R. C. O.; Yoshida, M. I.; Mussel, W. N.; Fialho, S. L.; Barbosa, J.; *J. Therm. Anal. Calorim.* **2017**, *127*, 577. [Crossref]
58. Ferreira, B. D. L.; Araújo, N. R. S.; Ligório, R. F.; Pujatti, F. J. P.; Mussel, W. N.; Yoshida, M. I.; Sebastião, R. C. O.; *J. Therm. Anal. Calorim.* **2018**, *134*, 773. [Crossref]
59. Ferreira, B. D. L.; Araújo, N. R. S.; Ligório, R. F.; Pujatti, F. J. P.; Yoshida, M. I.; Sebastião, R. C. O.; *Thermochim. Acta* **2018**, *666*, 116. [Crossref]
60. Koga, N.; Šesták, J.; *Thermochim. Acta* **1991**, *182*, 201. [Crossref]
61. Carvalho, F. S.; Braga, J. P.; Alves, M. O.; *J. Mol. Model.* **2022**, *28*, 286. [Crossref]
62. Rojek, B.; Suchacz, B.; Wesolowski, M.; *Thermochim. Acta* **2018**, *659*, 222. [Crossref]
63. Çepelioğullar, Ö.; Mutlu, İ.; Yaman, S.; Haykiri-Acma, H.; *J. Anal. Appl. Pyrolysis* **2016**, *122*, 84. [Crossref]
64. Monticeli, F. M.; Neves, R. M.; Ornaghi Jr., H. L.; *Cellulose* **2021**, *28*, 1961. [Crossref]
65. Araujo, N. R. S.; Sebastião, R. C. O.; Freitas-Marques, M. B.; da Nova Mussel, W.; Yoshida, M. I.; Virtuoso, L. S.; *Therm. Sci. Eng. Prog.* **2022**, *36*, 101490. [Crossref]
66. França, C. S.; Porto, D. L.; Junior, J. H. S.; Souza, F. S.; Nascimento, T. G.; Nogueira, F. H. A.; Gomes, A. P. B.; Jordão, A. K.; Ribeiro, E. H. S.; Araújo, N. R. S.; Sebastião, R. C. O.; Aragão, C. F. S.; *J. Therm. Anal. Calorim.* **2022**, *147*, 13395. [Crossref]
67. Queiroz, E. C.; Marques, L. F.; Carvalho, G. R.; Gomes, E. M.; Franco, D. F.; Quirino, W. G.; Rodrigues, J. S. M.; Freire, R. O.; Freitas-Marques, M. B.; Araújo, N. R. S.; Yoshida, M. I.; Sebastião, R. C. O.; Pinheiro, C. B.; Machado, L. O. M.; Machado, F. C.; *Opt. Mater.* **2024**, *154*, 115672. [Crossref]
68. de Freitas-Marques, M. B.; Araújo, N. R. S.; Rosa Jr., A. A.; Yoshida, M. I.; de Oliveira Sebastião, R. C.; da Nova Mussel, W.; *J. Therm. Anal. Calorim.* **2024**, *149*, 7329. [Crossref]
69. Araújo, N. R. S.; Amaral, L. V.; Pujatti, F. J. P.; Freitas-Marques, M. B.; Mussel, W. N.; Sebastião, R. C. O.; *J. Therm. Anal. Calorim.* **2022**, *147*, 14371. [Crossref]
70. Wakimoto, S.; Matsukawa, Y.; Numazawa, Y.; Matsushita, Y.; Aoki, H.; *Fuel* **2023**, *343*, 127836. [Crossref]
71. Asif, M. R.; Bording, T. S.; Barfod, A. S.; Grombacher, D. J.; Maurya, P. K.; Christiansen, A. V.; Auken, E.; Larsen, J. J.; *IEEE Access* **2021**, *9*, 34635. [Crossref]

72. Sbirrazzuoli, N.; Girault, Y.; Elegant, L.; *Die Angewandte Makromolekulare Chemie* **1993**, *211*, 195. [Crossref]
73. Cai, J.; Liu, R.; *Ind. Eng. Chem. Res.* **2009**, *48*, 3249. [Crossref]
74. Pérez-Maqueda, L. A.; Criado, J. M.; Gotor, F. J.; Málek, J.; *J. Phys. Chem. A* **2002**, *106*, 2862. [Crossref]
75. Sbirrazzuoli, N.; Girault, Y.; Elégant, L.; *Thermochim. Acta* **1995**, *260*, 147. [Crossref]
76. Araujo, B. C. R.; Carvalho, F. S.; Marques, M. B. F.; Braga, J. P.; Sebastião, R. C. O.; *J. Braz. Chem. Soc.* **2020**, *31*, 1392. [Crossref]

Submitted: November 25, 2024
Published online: May 19, 2025