



# ANN-MATOPT hybrid algorithm: determination of kinetic and non-kinetic parameters in different reaction mechanisms

M. M. Canedo Alonso<sup>1</sup> · Jaime González Cuadra<sup>1</sup> · J. L. González-Hernández<sup>1</sup>

Received: 29 April 2021 / Accepted: 23 July 2021  
© The Author(s) 2021

## Abstract

In this work we have applied the computational methodology based on Artificial Neural Networks (ANN) to the kinetic study of distinct reaction mechanisms to determine different types of parameters. Moreover, the problems of ambiguity or equivalence are analyzed in the set of parameters to determine in different kinetic systems when these parameters are from different natures. The ambiguity in the set of parameters show the possibility of existence of two possible set of parameter values that fit the experimental data. The *deterministic analysis* is applied to know beforehand if this problem occurs when rate constants of the different stages of the mechanism and the molar absorption coefficients of the species participating in the reaction are obtained together. Through the *deterministic analysis* we will analyze if a system is identifiable (unique solution or finite number of solutions) or if it is non-identifiable if it possesses infinite solutions. The determination of parameters of different nature can also present problems due to the different magnitude order, so we must analyze in each case the necessity to apply a second method to improve the values obtained through ANN. If necessary, an optimization mathematical method for improving the values of the parameters obtained with ANN will be used. The complete process, ANN and mathematical optimizations constitutes a hybrid algorithm ANN-MATOPT. The procedure will be applied first for the treatment of synthetic data with the purpose of checking the applicability of the method and after, it will be used in the case of experimental kinetic data.

**Keywords** Computational chemical kinetics · Artificial neural networks · Hybrid algorithm · Ambiguity or equivalence of parameters

---

✉ M. M. Canedo Alonso  
mcanedo@usal.es

<sup>1</sup> Department of Physical Chemistry, Faculty of Chemistry, University of Salamanca, Plaza de los Caídos s/n, 37008 Salamanca, Spain

## 1 Introduction

The application of computational techniques to the study of reaction kinetic systems has been very important since it has been possible to determine kinetic, analytic and/or thermodynamic parameters, allowing also, the discrimination between the different mechanisms that could be responsible of the course of the reaction.

The computational techniques used are numerous and are based in very diverse methodologies such as non-linear regression methods [1–5], curve resolution techniques [6, 7] or techniques that use methodologies based in Artificial Neural Networks [8–11].

In a previous work we have used the ANN methodology for the determination of rate constants from kinetic data in the case of a system of consecutive irreversible first-order reactions [11]. The ANN method was applied for the study of the reaction that takes place between 2-mercaptoethanol with Carbonyl cyanide 3-chlorophenylhydrazone, from experimental kinetic data the determination of the reaction rate constants was carried out, performing first an experimental design of the type *Central Star Composite Experimental Design* (CSCED) and then, the prediction of the rate constant through ANN.

In later works a hybrid algorithm, ANN-AGDC, was developed, consisting of the sequential application of two methods: Artificial Neural Network method (ANN) and mathematical optimization algorithm AGDC. This hybrid algorithm has been used for the treatment of kinetic data for the determination of rate constants and for the discrimination between reaction mechanisms [12]. The application of hybrid algorithm is a great advantage since the ANN, that is applied first, is a treatment that does not need initial estimates, so that the obtained results can be used as initial estimates for the second method (mathematical optimization) that does require the use of previous values for the parameters to determine.

The first version of hybrid algorithm ANN-AGDC was applied to determine Activation Thermodynamic Parameters (ATP) from the treatment of non-isothermal kinetic data coming from diverse reaction mechanisms [13–15]. In the first two works the applicability of hybrid algorithm ANN-AGDC is checked for the determinations of ATP from non-isothermal kinetic data, directly, without the need to determine the kinetic constants in the previous step. In these works, the ANN methodology was first applied and the need to apply the second process, the mathematical optimization algorithm AGDC, was checked. Subsequently, this methodology was used for the treatment of experimental non-isothermal kinetic data from the isomerization reaction of the steroid 5-cholesten-3-one to 4-cholesten-3-one catalyzed by sodium ethoxide and from the kinetics for the breakdown of the trinuclear chromium acetate cluster with a series of monoprotic and diprotic amino acid ligands [15].

In this work we have used a new hybrid algorithm ANN-MATOPT for the determination of rate constants and molar absorption coefficients simultaneously, in the case of different reactions mechanisms whose stages are of first order. The studied mechanisms are the following:

- Model I: Simple irreversible reaction in which a reactant gives a product through a single step.
- Model II: System formed by two irreversible reactions in which two reactants give the same product.
- Model III: Mechanisms in which the same reactant gives two different products through two irreversible reactions.
- Model IV: System of consecutive irreversible reactions.

When parameters of different nature are determined jointly and simultaneously, there can be problems of ambiguity or equivalence in the set of parameters to determine, which indicates the possible existence of two set of values of the parameters that fit the experimental data [4, 5, 16–18]. Through the *deterministic analysis* it can be known beforehand if a model is identifiable with a unique solution or with a finite number of solutions, or non-identifiable with infinite solutions [16]. First, by means of *deterministic analysis*, for each one of the mechanisms, we will assess if ambiguity occurs when the rate constants of the different stages of the mechanism and the molar absorption coefficients of the participating species in the reaction are jointly determined. Subsequently, we will check in the case of situations where there is ambiguity if the application of ANN leads us to a unique solution or to any of the mathematically acceptable solutions.

The problem of the ambiguity in the parameters does not have a solution from a mathematical point of view so we apply the ANN methodology, either individually or in conjunction with a mathematical optimization algorithm, for situations in which this problem is found and we will check if the obtained results lead us to an unique solution or to all of those mathematically accepted.

The determination of parameters of different nature can also present problems due to their different magnitude order, so we need to analyze in each case the need to apply a second method (mathematical optimization) to improve the values obtained through ANN.

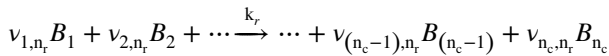
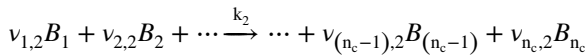
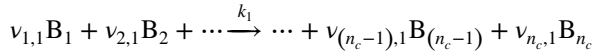
In the case of models I and II, whichever are the set of parameters to determine, exists an unique solution, so the models are identifiable. In model III depending on the set of parameters determined, the model can be identifiable with an unique solution or non-identifiable with infinite solutions.

In the case of model IV formed by two irreversible first order consecutive reactions, when the joint determination of rate constants and the molar absorption coefficient of the reactant or the final product is carried out, there exists a unique solution, so the system is identifiable in both cases. In case that the rate constants and the molar absorption coefficient of the intermediate species are jointly determined, an ambiguity problem occurs in the solution, finding two possible solutions that cannot be distinguished from the mathematical point of view [4, 5]. This aspect is clearly displayed in the case of the determination of these parameters through mathematical optimization procedures.

## 2 The oretical aspects

### 2.1 Chemical kinetics aspects

Each one of the  $n_r$  chemical elementary reaction that can occur in a chemical system in which a number of species  $n_c$  are present, can be written using the following equation [19]:



$n_c$  = number of chemical species involved in the mechanism.  $n_r$  = number of chemical elementary reaction of the mechanism.  $k_r$  = kinetic rate constant for each elementary reaction  $r$  ( $r = 1, \dots, n_r$ ).  $B_j$  = chemical species involved in the reaction mechanism ( $j = 1, \dots, n_c$ ).  $\nu_{j,r}$  = stoichiometric coefficient of the specie  $B_j$  in the reaction  $r$  ( $\nu_{j,r} < 0$  for the reagents and  $\nu_{j,r} > 0$  for the products).

The elementary reaction  $r$  occurs at a rate  $v_r$ :

$$\begin{aligned} r = 1 \quad v_1 &= k_1 [B_1]^{|\nu_{1,1}|} \cdot [B_2]^{|\nu_{2,1}|} \dots [B_c]^{|\nu_{c,1}|} = k_1 \prod_{c=1}^{n_c} [B_c]^{|\nu_{c,1}|} \\ r = 2 \quad v_2 &= k_2 [B_1]^{|\nu_{1,2}|} \cdot [B_2]^{|\nu_{2,2}|} \dots [B_c]^{|\nu_{c,2}|} = k_2 \prod_{c=1}^{n_c} [B_c]^{|\nu_{c,2}|} \\ &\dots \\ r = r \quad v_r &= k_r [B_1]^{|\nu_{1,r}|} \cdot [B_2]^{|\nu_{2,r}|} \dots [B_c]^{|\nu_{c,r}|} = k_r \prod_{c=1}^{n_c} [B_c]^{|\nu_{c,r}|} \end{aligned} \quad (1)$$

where  $B_c$  are the species that act as reagents in the elementary reaction  $r$ ,  $\nu_{c,r}$  is the stoichiometric coefficient of the species  $B_c$  in the reaction  $r$  ( $\nu_{c,r} < 0$  for reagents),  $[B_c]$  is the concentration of the species  $B_c$  and  $k_r$  is the kinetic rate constant for each elementary reaction  $r$ .

The variation of the concentration of species  $B_j$  in each of the  $r$  elementary reactions, in a time interval, is equal to:

$$\begin{aligned} r = 1 \quad \frac{d[B_j]}{dt} &= \nu_{j,1}v_1 \\ r = 2 \quad \frac{d[B_j]}{dt} &= \nu_{j,2}v_2 \\ &\dots \\ r = r \quad \frac{d[B_j]}{dt} &= \nu_{j,r}v_r \end{aligned} \quad (2)$$

The total variation of the concentration is expressed through the following differential equation:

$$\frac{d[B_j]}{dt} = \sum_{r=1}^{n_r} \nu_{j,r} \nu_r = \sum_{r=1}^{n_r} \nu_{j,r} k_r \prod_{l=1}^{n_c} [B_c]^{|\nu_{c,r}|} \quad (3)$$

The change in the concentration of each of the  $n_c$  species involved in the mechanism is given by  $d[B_j]/dt$ , so we will have a system of differential rate equations whose resolution provides the concentration of each specie in the considered time interval,  $([B_j]_{t_i})$ .

The reaction mechanisms studied in this work are the following:

1. Model I  $B_1 \xrightarrow{k} B_2$
2. Model II  $B_2 \xrightarrow{k_2} B_1$   $B_3 \xrightarrow{k_3} B_1$
3. Model III  $B_1 \xrightarrow{k_2} B_2$   $B_1 \xrightarrow{k_3} B_3$
4. Model IV  $B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$

Table 1 shows the systems of differential rate equations obtained for each one of the mechanisms and the resulting expressions for the concentration of each species obtained from the resolution of the system of differential rate equations.

If the development of the chemical reaction is followed spectrophotometrically measuring the total absorbance of the reacting mixture and in the event that the optical path length  $l = 1$  cm, according to the Lambert–Beer–Bouguer's Law at time  $i$  and at wavelength  $\lambda$ , we obtain:

**Table 1** Reaction mechanism, differential rate equations and concentration/time equations for each reaction mechanism

Model	Differential rate equations	Concentration/time equations
$B_1 \xrightarrow{k} B_2$	$\frac{d[B_1]}{dt} = -k[B_1]$	$[B_1] = [B_1]_0 e^{-kt}$
	$\frac{d[B_2]}{dt} = k[B_1]$	$[B_2] = [B_1]_0 (1 - e^{-kt})$
$B_2 \xrightarrow{k_2} B_1$	$\frac{d[B_2]}{dt} = -k_2[B_2]$	$[B_2] = [B_2]_0 e^{-k_2 t}$
$B_3 \xrightarrow{k_3} B_1$	$\frac{d[B_3]}{dt} = -k_3[B_3]$	$[B_3] = [B_3]_0 e^{-k_3 t}$
	$\frac{d[B_1]}{dt} = k_2[B_2] + k_3[B_3]$	$[B_1] = [B_2]_0 + [B_3]_0 - [B_2]_0 e^{-k_2 t} - [B_3]_0 e^{-k_3 t}$
$B_1 \xrightarrow{k_2} B_2$	$\frac{d[B_2]}{dt} = k_2[B_1]$	$[B_1] = [B_1]_0 e^{-(k_2+k_3)t}$
$B_1 \xrightarrow{k_3} B_3$	$\frac{d[B_3]}{dt} = k_3[B_1]$	$[B_2] = \frac{k_2[B_1]_0}{(k_2+k_3)} (1 - e^{-(k_2+k_3)t})$
	$\frac{d[B_1]}{dt} = -(k_2+k_3)[B_1]$	$[B_3] = \frac{k_3[B_1]_0}{(k_2+k_3)} (1 - e^{-(k_2+k_3)t})$
$B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$	$\frac{d[B_1]}{dt} = -k_1[B_1]$	$[B_1] = [B_1]_0 e^{-k_1 t}$
	$\frac{d[B_2]}{dt} = k_1[B_1] - k_2[B_2]$	$[B_2] = \frac{k_1[B_1]_0}{k_2-k_1} (e^{-k_1 t} - e^{-k_2 t})$
	$\frac{d[B_3]}{dt} = k_2[B_2]$	$[B_3] = [B_1]_0 \left( 1 - \frac{k_2}{k_2-k_1} e^{-k_1 t} + \frac{k_1}{k_2-k_1} e^{-k_2 t} \right)$

$$A_{i,\lambda} = \sum_{j=1}^{n_c} \epsilon_{j,\lambda} x [B_j]_i \quad (4)$$

where  $n_c$  is the number of species that absorb at the wavelength  $\lambda$ ,  $A_{i,\lambda}$  is the total absorbance of the reacting mixture at time  $i$  and at wavelength  $\lambda$ ,  $\epsilon_{j,\lambda}$  is the molar absorption coefficient of the specie  $j$  at wavelength  $\lambda$  and  $[B_j]_i$  the molar concentration of the specie  $j$  at time  $i$ .

## 2.2 Hybrid algorithm

A hybrid algorithm is a combination of different techniques that solve the same problem [20]. The hybrid algorithm used in this work is formed by a combination of two algorithms based in different mathematical principles and they are applied sequentially. In the first stage the Artificial Neural Network (ANN) method is applied for the determination of the individual rate constants and the molar absorption coefficients. The values of the parameters obtained through the application of ANN are used as initial estimates of a mathematical optimization algorithm applied in the second stage with the objective of improving the final values of the parameters.

### 2.2.1 Artificial neural networks (ANN)

The Artificial Neural Network computational method, ANN, is a systematic procedure of data processing formed by a large number of simple elements (nodes or artificial neurons) interconnected with each other [21–23]. Each neuron receives multiple input signals and sends a single output signal for which it contains two algorithms [23], one of them calculates the weighted sum of the values that are received by the input connections ( $x_i$ ) and the other one generates a response or output ( $x_j$ ) that is transferred to other neurons. The neurons network learns through the adjustment of the weights ( $w_{ji}$ ) of the connections between neurons until it provides predictions with enough precision.

Each neuron determines the net input value,  $S_j$ , from the weighted sum of the input values  $\left( S_j = \sum_i x_i w_{ji} \right)$ , later it is transformed into the activation value,  $a_j$ ,  $a_j(k) = F_i(a_j(k-1), S_j(k))$ , where  $a_j(k)$  is the actual activation value,  $a_j(k-1)$  the previous activation value and  $S_j(k)$  the net input value.

The output value is determined by applying the output function (*transfer function*), in most cases the activation and the net input are identical, so  $x_j = f_j(a_j) = f_j(S_j)$ .

In an ANN the neurons are connected and organized in levels or layers (architecture);

- *Input layer* where data is presented to the network.
- *Output layer* where the network response is provided.
- *Hidden layer* formed by those neurons whose inputs come from previous layers and whose outputs pass to neurons from subsequent layers.

The simplest network (monolayer network) consists of a group of organized neurons in one layer, between which lateral, crossed and recurrent connections are established. The multilayer networks consist of clustered sets of neurons in various levels or layers where the output of one layer is the input of the next one. The neurons of one layer receive input signals from a previous layer and send output signals to a following layer (feedforward) or the output of the neurons from posterior layers can be connected to the input of preceding layers, (feedback).

From a data set, training data set, the ANN learns to calculate the correct output for each input (training), during this process, the weights of the ANN connections suffer modifications in response to the input information ( $w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji}(k)$ ). The ANN has learnt when the values of the weights remain stable ( $dw_{ji}/dt = 0$ ).

The goal of the training of an ANN is achieving that for a set of inputs the desired or consistent set of outputs is produced.

### 2.2.2 Mathematical optimization

The mathematical optimization methods usually use the criteria of least squares for the determination of parameters from experimental data. The procedure consists of minimizing the error expressed as the sum of the squares of the differences between the values of the experimentally observed magnitude and those calculated using the equation of the proposed mechanism or model (sum of the quadratic deviation function, SQD) [24, 25]. In the case that the observed magnitude is the total absorbance of the reacting mixture, the SQD function has the form:

$$\text{SQD}(\mathbf{X}) = \sum_{i=1}^{n_d} ((A_i)_C(\mathbf{t}, \mathbf{X}) - (A_i)_E)^2 \quad i = (1, 2 \dots n_d) \quad (5)$$

$n_d$  = number of pairs of the absorbance/time data.  $(A_i)_E$  = experimental absorbance data.  $(A_i)_C$  = calculated absorbance data.  $\mathbf{t}$  = values for the independent variable (time).  $\mathbf{X}$  = vector that contains the parameters to be determined ( $X_1, X_2, \dots, X_p$ ). It is possible to fit the function  $y = (A_i)_C(\mathbf{t}, \mathbf{X})$  to the experimental data  $(A_i)_E$ , minimizing the SQD function (mathematical optimization) and adjusting the parameters  $X_1, X_2, \dots, X_p$  to obtain the minimum of SQD [24, 25]. The mathematical optimization is an iterative process so in order to initiate the process, it is necessary a first estimate of the values of the parameter,  $\mathbf{X}^{(0)}$ , and with them the first value of the function SQD,  $\text{SQD}^{(0)}$ , will be obtained. When the optimization is finished the values of the parameters,  $\mathbf{X}^*$ , that provide the smallest value of SQD,  $\text{SQD}^*$ , will be obtained.

Gradient methods (Gauss–Newton, Levenberg–Marquardt, Davidon–Fletcher–Powell, etc., ...) are often used to carry out the optimization, these methods are based on the Taylor series development of the SQD function [24, 25].

$$\text{SQD}(\mathbf{X} + \Delta\mathbf{X}) \approx \text{SQD}(\mathbf{X}) + \mathbf{g}^T \Delta\mathbf{X} + \frac{1}{2} \Delta\mathbf{X}^T \mathbf{H} \Delta\mathbf{X} + \dots \quad (6)$$

$\mathbf{X}$  = vector that contains the parameters to be determined.  $\Delta\mathbf{X}$  = vector constituted by the variations of the parameter values.  $\mathbf{g}$  = Jacobian gradient vector of the SQD

function ( $\mathbf{g}^T$  = transposed vector) whose components are the derivatives of the SQD function with respect to one of the parameters  $X_p$ .  $\mathbf{H}$  = Hessian matrix of the SQD function (symmetric matrix  $n_p \times n_p$  containing the second-order partial derivatives of the SQD function).

In Eq. (5) we can dispense the terms in which the derivatives higher than the first order (first order gradient methods) or the terms in which the derivatives higher than the second order intervene (second order gradient methods). From the calculation of the derivative of the SQD function with respect to each of the parameters, its cancellation and the resolution of the equations obtained, the point in which the function SQD has a minimum value ( $\mathbf{X}^*$ ) is reached, Eqs. (7) and (8) for first and second order gradient methods respectively:

$$\mathbf{X}^* \approx \mathbf{X} - \mathbf{g}(\mathbf{X}) \quad (7)$$

$$\mathbf{X}^* \approx \mathbf{X} - \mathbf{H}(\mathbf{X})^{-1} \mathbf{g}(\mathbf{X}) \quad (8)$$

The equation that provides the value  $\mathbf{X}^*$  can be used to move iteratively towards the SQD minimum starting from an initial estimate  $\mathbf{X}^{(0)}$  and generating values  $\mathbf{X}^{(m)}$  at each iteration, so that the SQD value is less than from the previous iteration.

$$\mathbf{X}^{(m+1)} = \mathbf{X}^{(m)} - \mathbf{g}(\mathbf{X}^{(m)}) \quad (9)$$

$$\mathbf{X}^{(m+1)} = \mathbf{X}^{(m)} - \mathbf{H}(\mathbf{X}^{(m)})^{-1} \mathbf{g}(\mathbf{X}^{(m)}) \quad (10)$$

The movement vector will be,  $\Delta \mathbf{X}^{(m)} = -\mathbf{g}(\mathbf{X}^{(m)})$  for first order methods and  $\Delta \mathbf{X}^{(m)} = -\mathbf{H}(\mathbf{X}^{(m)})^{-1} \mathbf{g}(\mathbf{X}^{(m)})$  for the second ones, where  $m=0, 1, \dots$  is the corresponding iteration.

In each iteration the vector  $\Delta \mathbf{X}^{(m)}$  must provide a point  $\mathbf{X}^{(m+1)}$  that gives rise to a lower value of the SQD function, avoiding divergence. The first order method always evolves towards the minimum, but it is excessively slow. Second-order methods, due to the approximations that are imposed and to the not strictly quadratic condition of the SQD function, present problems that can lead to the divergence of the process. There are modifications of the Gauss–Newton method that improve the results: Levenberg–Mardquardt method that modifies the vector  $\mathbf{X}^{(m+1)}$  by introducing a parameter  $\lambda$ , Davidon–Fletcher–Powell method that determines an inverse matrix approximated by successive approximations or the AGDC that introduces control devices to avoid divergence [3–5, 24, 25].

The hybrid algorithm used in this work uses the values of the parameters obtained with ANN as initial estimates ( $\mathbf{X}^{(0)}$ ) for the process of mathematical optimization (MATOPT).

### 2.3 Ambiguity or equivalence in the parameters

When the treatment of kinetic data corresponding to a certain reaction mechanism is carried out in order to determine a set of parameters, there might be one or more sets



of parameters ( $\mathbf{X}=(X_1, \dots, X_p)$ ,  $\mathbf{X}'=(X'_1, \dots, X'_p)$ , etc,..) that fit the experimental data: ambiguity or equivalence in the set of parameters [4, 5, 16–18]. This problem has been studied through different methods, one of them consists of the application of the *deterministic analysis* [16] that allows to detect it beforehand and determine the relationship between the solutions  $\mathbf{X}$  and  $\mathbf{X}'$ .

In a chemical reaction whose development is experimentally followed measuring the total absorbance, it might happen that the molar absorption coefficients of some species involved in the reaction ( $\varepsilon_{B_j}$ ) are unknown parameters and it is necessary to determine them jointly with the rate constants,  $\mathbf{X}=(k_1, \dots, k_r, \varepsilon_{B_1}, \dots, \varepsilon_{B_j})$ .

In the case of reaction mechanism of first order, we consider the following vectors and the corresponding components:  $\mathbf{x}_0(\mathbf{X})$ , initial concentrations;  $\mathbf{x}(t, \mathbf{X})$ , concentrations;  $\mathbf{x}'(t, \mathbf{X})$ , rate differential equations and  $\mathbf{y}(t, \mathbf{X})$ , absorbance values (linear function that depends of the parameters  $\mathbf{X}$  and the concentrations). According to the *deterministic analysis* two sets of the parameters values,  $\mathbf{X}$  and  $\mathbf{X}'$  ( $\mathbf{X} \neq \mathbf{X}'$ ) are not differentiable among them if:

$$\mathbf{y}(t, \mathbf{X}') = \mathbf{y}(t, \mathbf{X}) \quad (11)$$

According to this equation, the following situations can be given:

- If  $\mathbf{X}' = \mathbf{X}$ , the solution is unique, the model is *only identifiable* in  $\mathbf{X}$ .
- If there is a *finite number* of distinct solutions,  $\mathbf{X}' \neq \mathbf{X}$ , the model is *identifiable* in  $\mathbf{X}$ , but *the solution is not unique*.
- If there is an *infinite number* of solutions, the model is *non-identifiable*.

The approximation that applies the Laplace transform [16] to the differential rate equations and to the equations that relates the absorbance and the concentration, is one of the most used mathematical methods to analyze these problems. Through this method we obtain the transform of the concentration vector,  $Z(s, \mathbf{X})$  ( $s$  is a complex argument) and the Laplace transform of the absorbance,  $Y(s, \mathbf{X})$ . Each component  $Y_i(s, \mathbf{X})$  of  $Y(s, \mathbf{X})$  is a rational function with the form:

$$Y_i(s, \mathbf{X}) = \frac{\phi_{n+1}^i s^{n-1} + \dots + \phi_{2n}^i}{s^n + \phi_1^i s^{n-1} + \dots + \phi_n^i} \quad (12)$$

The coefficients  $\phi_j^i$  depend on  $\mathbf{X}$  and  $x_0^{(i)}$  and the vector  $\Phi$  will be formed by all the coefficients  $Y_1(s, \mathbf{X}) \dots Y_m(s, \mathbf{X})$ .

Considering another set of parameters  $\mathbf{X}' = (k'_1, \dots, k'_r, \varepsilon'_{B_1}, \dots, \varepsilon'_{B_j})$  whose Laplace transform is  $Y(s, \mathbf{X}')$ , the Eq. (10) is satisfied if and only if:  $Y(s, \mathbf{X}') = Y(s, \mathbf{X})$ . This relationship is true if and only if  $\Phi(\mathbf{X}') = \Phi(\mathbf{X})$ , from which a system of polynomial equations is obtained, establishing the number of solutions of the system of equations we can establish if a reaction model possess more than one solution, meaning that it is identifiable.

If  $n_p$  is the number of parameters to optimize and  $n_e$  is the number of polynomial equations that can be established, the following situations can occur:

- $n_p < n_e$  then the system is identifiable and the solution is unique.
- $n_p = n_e$  then the system is identifiable and the solution is not unique.
- $n_p > n_e$  then the system is not identifiable, it has infinite solutions.

### 3 Computational aspects

#### 3.1 ANN computational application

In this work, the applications provided by MATLAB® [26] are used for the design and training of the ANN and subsequent prediction of the parameters. All the applications are available in the *Neural Network Toolbox*™, since a set of parameters have been determined from kinetic data, the application that allows to adjust functions, *Fitting a Function*, has been used. In order to use this application, it is necessary to provide to the ANN a data set (training data set) to use them in the training of the ANN, this set needs to be formed by a *inputs* matrix that contains data corresponding to different kinetic experiences and a matrix (*targets*) consisting of the parameters corresponding to the data of the *inputs* matrix. The data is generated synthetically for each specific case and in the training the ANN learns to establish relationships between the *inputs* and *targets* matrixes.

##### 3.1.1 Generation of data

The synthetic total absorbance/time data for the training of the ANN are generated computationally through the executable programs (type “.m”) of MATLAB, considering the initial concentration conditions, the time interval in which the reaction takes place, the rate constant values ( $k_r$ ) and the molar absorption coefficients ( $\epsilon_{B_j}$ ) obtained from the corresponding experimental desing. The concentration of each one of the chemical species in the time interval considered, is determined from the expressions that result from the resolution of the system of rate differential equations (Table 1) and the values of the absorbance from the Lambert-Beer-Bouguer’s law (Eq. 3). The absorbance data generated will be distributed in the *inputs* matrix and the values of the parameters with which they have been generated (constants and molar absorption coefficients) in the *targets* matrix. The dimensions of the *inputs* and *targets* matrixes, number of rows  $\times$  number of columns, are the following:

- *Inputs*: number of rows = number of data from each experience, number of columns = number of kinetic curves.
- *Targets*: number of rows = number of parameters, number of columns = number of kinetic curves.

### 3.1.2 Design and training of the ANN

In each case studied it is necessary to find the optimal architecture of the ANN, that is, the optimal number of layers and nodes of each one of the hidden layers. The procedure to perform is the subsequent:

- Desing of an ANN formed by  $n$  layers ( $(n - 1)$  *Hidden* + 1 *Output*):
- Establish the number of layers ( $numLayer = n$ ).
- Establish the number of neurons in the hidden layers ( $numHiddenNeurons1 = Integer\ number, numHiddenNeurons2 = Integer\ number, \dots$ )
- Create the network ( $net = newfit(inputs, targets, [numHiddenNeurons1, \dots])$ )
- Training of the ANN:  $net = train(net, inputs, targets)$

The training of each one of the ANNs designed, is carried out with the data corresponding to the training data set. These data is divided into three groups: experiences for the training (TR), experiences to validate the network (VL) and those that are used as an independent test (TS). The program automatically establishes the percentages for the curves for TR, VL and TS, but can be modified by the used to improve the results. MATLAB applies by default the Levenberg–Marquardt algorithm to carry out the training, although it is possible to choose another algorithm among others provided by MATLAB.

The result of the training can be considered acceptable if a series of requisites are met: the mean square errors (MSE) is significantly small, the VL and TS errors have similar characteristics, there is no overfitting and the linear regression between the corresponding *outputs* and *targets* for TR, VL and TS provide values for the correlation coefficient,  $R$ , close to 1.

If the training of the ANN has been carried out correctly, it could be considered suitable to carry out prediction tasks, otherwise we can improve the results: reestablish the initial ANN and train again (retrain), increase the number of neurons in the hidden layers, increase the number of hidden layers, modify the percentage of experiences for TR, VL and TS or use another optimization algorithm that minimizes the difference of the values of the *output* and *targets*.

### 3.1.3 Prediction

The optimal ANN can be used to determine unknown parameters (prediction), carrying out the following steps:

1. Provide data to the ANN of the experiences whose parameters are unknown, *inputs*:  $N^\circ$  of rows =  $N^\circ$  data of each experience (it should match with the  $n^\circ$  of rows of the *inputs* matrix with which the ANN has been trained) and  $N^\circ$  columns =  $N^\circ$  of experiences whose parameters are unknown.
2. Prediction with the optimal ANN previously trained (*net*) through the command:  $[Y] = sim(net, inputs)$  ( $Y$  = values of the parameters predicted by the ANN).

### 3.2 Mathematical optimization

To carry out the mathematical optimization process the tool *Optimization Toolbox™* from MATLAB® has been used. Among all the possibilities that MATLAB offers, we have used the *lsqcurvefit* function, that used different gradient methods to perform the optimization process, Levenberg–Marquardt, Gauss–Newton, etc., ... that can be chosen by the user.

From the input data, *xdata* and *ydata*, the function *lsqcurvefit* finds the coefficients *x* that best fit the equation  $F(x, xdata)$ :

$$\min x \frac{1}{2} F(x, xdata) - ydata^2 = \frac{1}{2} \sum_i (F(x, xdata_i) - ydata_i)^2 \quad (13)$$

where *xdata* is the vector that contains the time value considered ( $t_i$ ), *ydata* is the vector that contains the absorbance experimental data and  $F(x, xdata)$  is the vector that contains the values calculated for the property measured (absorbance).

The sequence of commands is the following:

- Initial estimations of the parameters (values obtained with ANN):  $\times 0 = [Par1, Par2, \dots]$
- Options for the optimization according to the manual specifications: `options = optimset('MaxFunEvals', 1e8, 'TolFun', 1e-18, 'TolX', 1e-12, 'MaxIter', 1e4)`
- Order to carry out the optimization: `[x, resnorm] = lsqcurvefit(function, x0, xdata, ydata, 0, 1000, options)`

*x* = vector that contains the optimized parameters; *resnorm* = values of the sum of the quadratic deviations; *function* = name of the subprogram that needs to be created to calculate the absorbance data with the values estimated initially for the parameters and the ones that are been determined through the optimization process.

## 4 Results and discussion

In this work a hybrid algorithm (ANN-MATOPT) is applied to different reaction mechanisms to jointly determine parameters of different nature, specifically, rate constants ( $k_r$ ) and molar absorption coefficients ( $\epsilon_j$ ). This hybrid algorithm consists of a first stage where the ANN methodology is used to predict the values of  $k_r$  and  $\epsilon_j$  and a second stage where a mathematical optimization method is applied to the results obtained through ANN to improve them. In each reaction mechanism the *deterministic analysis* is applied to learn beforehand the existence of ambiguity and equivalence in the set of parameters to determine and it is checked if the hybrid algorithm ANN-MATOPT leads us to the true solution or the different set of parameters that fit the experimental data are obtained.

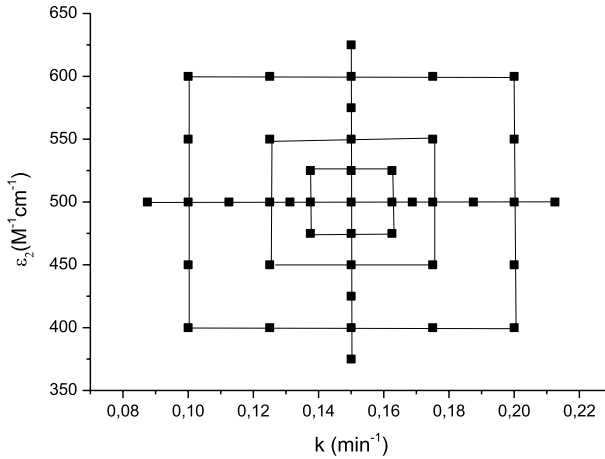


Fig. 1 Two-dimensional experimental design (2 factors,  $k_1$  and  $\epsilon_2$ ) for the system  $B_1 \rightarrow B_2$

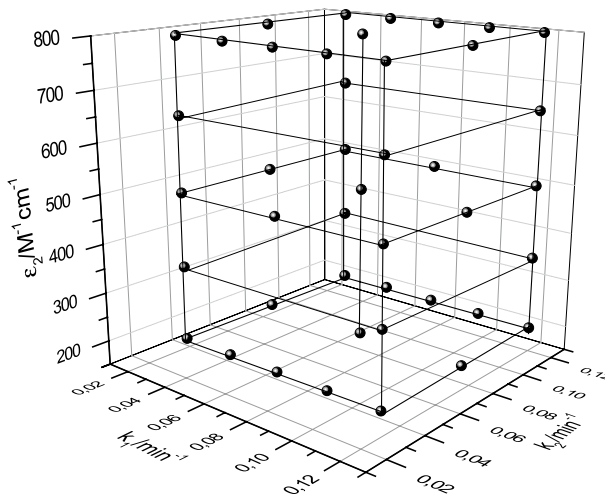


Fig. 2 Three-dimensional experimental design (3 factors  $k_1$ ,  $k_2$  and  $\epsilon_2$ ) for the system  $B_1 \rightarrow B_2 \rightarrow B_3$

#### 4.1 Experimental desing

The absorbance data set with which it is performed the ANN training consists of a set of total absorbance/time kinetic data (*inputs*) generated with the group (*targets*) of constants ( $k_r$ ) and the molar absorption coefficients ( $\epsilon_j$ ), organized according to the corresponding experimental design (ED) [27]. The ED used can be two or three dimensional depending on the number of *factors* involved, an example for each case is shown in Figs. 1 and 2.

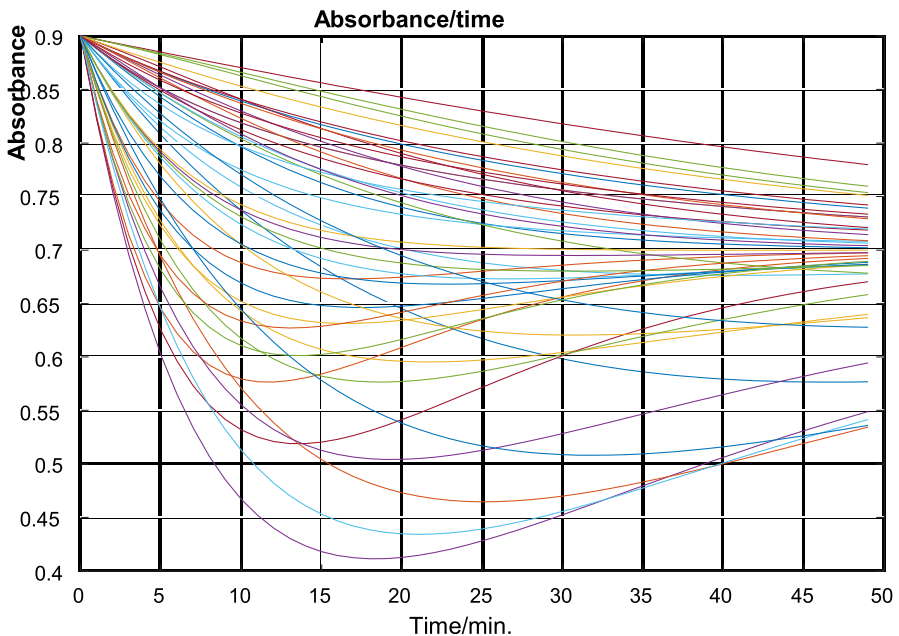
## 4.2 ANN training: optimal architecture

Once the kinetic data of total absorbance/time have been generated (*input* matrix) according to the ED (*targets* matrix), Fig. 3, we carry out the ANN training in order to find the most suitable to determine the parameters,  $k_r$ ,  $y$ ,  $\varepsilon_j$ , in each case. The set of data are divided in three groups to carry out the tasks of training, validation and testing (TR/TS/VL) in all the cases studied in this work, the optimal percentage is 80/10/10.

The optimal architecture of the ANN is determined using the *trial and error* method that consists of designing different ANNs with different number of hidden layers and different number of neurons in each one of them to carry out the training process for each ANN. The training is an iterative process in which in each iteration new values for the parameters are obtained (*outputs*) and it is controlled through the validation process, for which the program calculates the value of the mean square error (MSE) as follows:

$$MSE = \left( \frac{\sum_{i=1}^{n_p} \sum_{j=1}^{n_c} (b_{ij}^{output} - b_{ij}^{target})^2}{n_p \cdot n_c} \right)^{1/2} \quad (14)$$

where  $n_c$  is the number of kinetic curves,  $n_p$  number of parameters,  $b_{ij}^{output}$  and  $b_{ij}^{target}$  are the components of the *outputs* and *targets* matrixes.



**Fig. 3** 43 kinetic curves total absorbance/time for the system  $B_1 \rightarrow B_2 \rightarrow B_3$  generated from the corresponding experimental design

The program selects the minimum MSE value of VL reached (Fig. 4) that corresponds to the MSE value of the overall training process. The TS process is similar to the VL, it compares the elements of the *outputs* and *targets* matrixes in each time unit for the set of curves selected for the processes of TS and VL. The TS and VL curves must always run very close (Fig. 4), this should be a criterion to consider the correct training process.

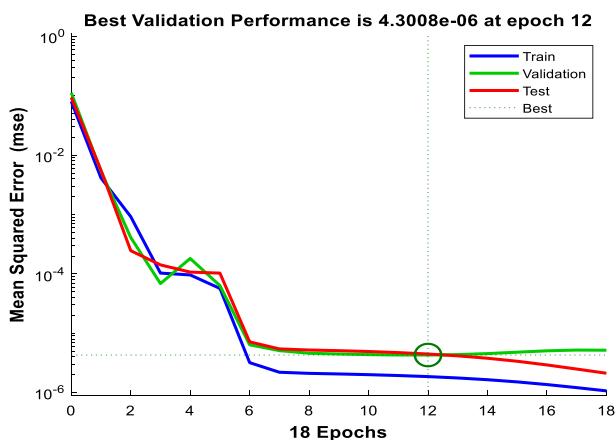
Once the ANN training process has been completed, the minimum MSE value and the values of the regression parameters of the overall training process and those of the TR, VL and TS groups are obtained. It can be considered that the training has been carried out correctly if the MSE value is small and if in the *outputs/targets* linear regressions (Fig. 5) the equations are linear (slope  $\approx 1$ , intercept  $\approx 0$ ) and the regression coefficient is  $R \approx 1$ .

When the optimal network is selected, the final results of the training are analyzed and it is verified that overfitting does not occur, verifying that the performance of the training and test set is good, that is, that the results obtained in the training process are acceptable in both cases (Table 2).

### 4.3 Prediction of parameters in different reaction mechanisms

In all the reaction mechanisms studied we have considered that the different stages are of order 1 and that the initial concentration ( $[B_j]_0$ ) of the reagents is different from 0 and that of the products is null. Besides,  $\mathbf{X} = (k_1, \dots, k_r, \varepsilon_{B_1}, \dots, \varepsilon_{B_j})$  and  $\mathbf{X}' = (k'_1, \dots, k'_r, \varepsilon'_{B_1}, \dots, \varepsilon'_{B_j})$  are the sets of parameters that fit the experimental data, being  $Y(s, \mathbf{X})$  and  $Y(s, \mathbf{X}')$  the Laplace transforms of the absorbance for  $\mathbf{X}$  and  $\mathbf{X}'$ .

In all of the cases, we proceed as follow:



**Fig. 4** Mean Square error (MSE) value in each iteration (epoch) for the training (TR), validation (VL) and testing (TS) processes

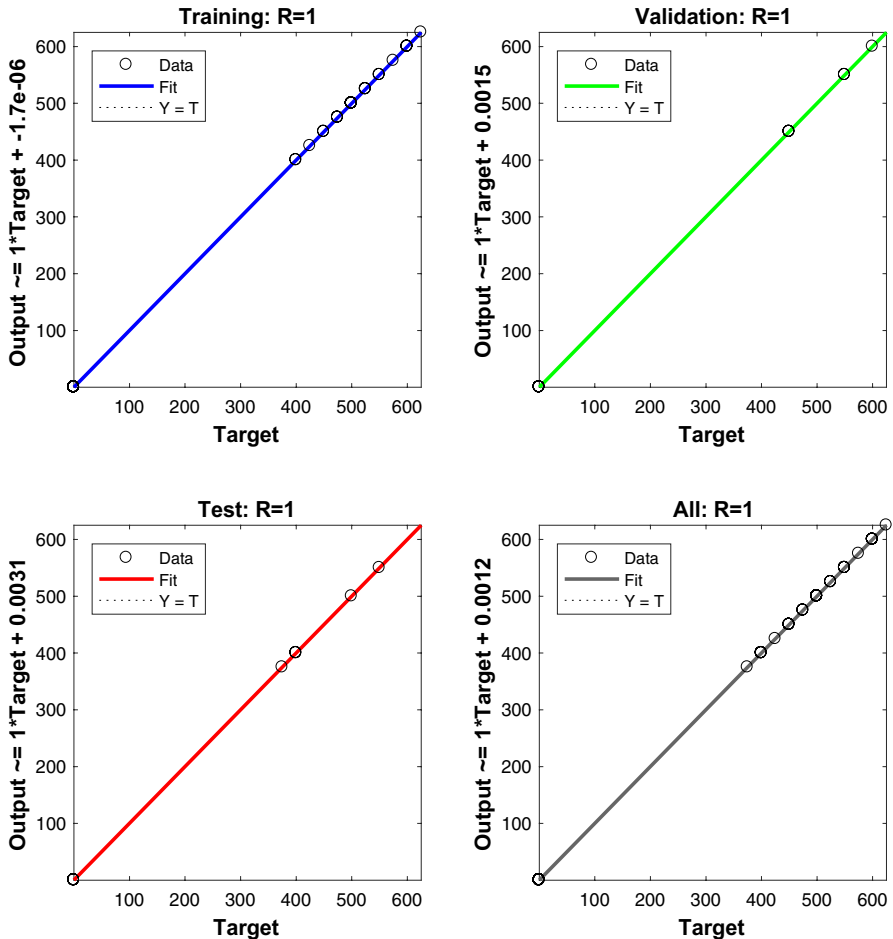


Fig. 5 Plots of the linear regression for the *outputs vs targets* matrixes for the training (TR), validation (VL), testing (TS) and total processes

- Establishment and resolution of the system of differential rate equations.
- Calculation of the Laplace transforms of the absorbance for  $\mathbf{X}$  and  $\mathbf{X}'$  ( $Y(s, \mathbf{X})$  and  $Y(s, \mathbf{X}')$ ).
  - $Y(s, \mathbf{X}) = Y(s, \mathbf{X}') \rightarrow$  We obtain a system of polynomial equations.
  - Solving the system of polynomial equations  $\rightarrow$  Knowledge beforehand of the *identifiability* of the system.
  - Experimental design in each case studied and generation of the  $A_T/t$  data for training the Artificial Neural Network (ANN).
  - Desing of the ANNs.
  - Training of the ANNs and analysis of the results  $\rightarrow$  Optimal ANN
  - Prediction of the parameters in experiences in which their value is not known



- Verification of the agreement of the results with that predicted by the *deterministic analysis* → *Identifiable system* with a unique solution or with a finite number of solutions, *non-identifiable system* with an infinite number of solutions.
- Verification of the need to apply a mathematical optimization method to improve the results → Hybrid algorithm ANN-MATOPT.

#### 4.3.1 Model $B_1 \xrightarrow{k} B_2$

For the case of this simple mechanism, the determination of the following groups of parameters are analyzed:  $\mathbf{X} = (k, \varepsilon_2)$ ,  $\mathbf{X} = (k, \varepsilon_1)$  and  $\mathbf{X} = (k, \varepsilon_1, \varepsilon_2)$ . The Laplace transform of the total absorbance for this case is expressed as:

$$Y(s, \mathbf{X}) = \frac{(s \varepsilon_{B_1} + \varepsilon_{B_2} k_1) [B_1]_0}{s^2 + sk_1} \quad (15)$$

For another set  $\mathbf{X}'$  parameters, the Laplace transform is  $Y(s, \mathbf{X}')$ , according to the *deterministic analysis* ( $Y(s, \mathbf{X}) = Y(s, \mathbf{X}')$ ) the following relationships are obtained:  $\varepsilon_{B_1} = \varepsilon'_{B_1}$ ,  $\varepsilon_{B_2} = \varepsilon'_{B_2}$  and  $k_1 = k'_1$ . Thus, the system is identifiable and the solution is unique in any of the cases. The ANN application for the prediction of the different group of parameters support the conclusions from the *deterministic analysis*, since in all of the cases the system is identifiable and the solution is unique.

Consider for example the prediction of ( $\mathbf{X} = k, \varepsilon_2$ ), after performing the corresponding ED, the training is carried out obtaining that for this case the optimal ANN consists of one hidden layer with 4 neurons. Then, a series of  $A_T/t$  curves are generated with certain values of  $k$  and  $\varepsilon_2$  that later we will assume that they are unknown and whose values are determined with the optimal ANN. In all the cases a solution is obtained and it is also observed the difference between the values of the parameters with which they have been generated and those calculated is minimal, so, it is no necessary to subsequently apply a mathematical optimization process.

**Table 2** Real values of  $k$  and  $\varepsilon_2$  for the system  $B_1 \rightarrow B_2$  and results obtained for the prediction through ANN

Exp	Real values		ANN calculated values		
	$k/\text{min}^{-1}$	$\varepsilon_2/\text{M}^{-1}\text{cm}^{-1}$	$k/\text{min}^{-1}$	$\varepsilon_2/\text{M}^{-1}\text{cm}^{-1}$	MSE
1	0.1600	550.0000	0.1600	550.0012	7.4578E-04
2	0.1200	500.0000	0.1206	499.9986	6.5465E-04
3	0.1500	450.0000	0.1475	450.0008	8.3476E-04
4	0.1375	475.0000	0.1416	475.0080	6.2465E-04
5	0.1100	525.0000	0.1406	530.0056	6.4578E-04
6	0.1400	530.0000	0.1406	530.0005	8.3468E-04
7	0.1100	600.0000	0.1180	599.9989	7.4573E-04
8	0.1800	520.0000	0.1789	519.9979	5.6749E-04
9	0.1125	500.0000	0.1150	499.9998	6.5454E-04
10	0.1650	380.0000	0.1653	380.0137	9.2476E-04

### 4.3.2 Model II $B_2 \xrightarrow{k_2} B_1 B_3 \xrightarrow{k_3} B_1$

For this mechanism, we study the determination of the following set of parameters:  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_3})$ . In this case the Laplace transform of the total absorbance is:

$$Y(s, \mathbf{X}) = \frac{\alpha[B_1]_0 + \beta[B_2]_0}{s^3 + s^2(k_2 + k_3) + sk_2k_3}$$

$$\alpha = (\varepsilon_{B_1}s^2 + s(\varepsilon_{B_1}k_2 + \varepsilon_{B_2}k_3) + \varepsilon_{B_1}k_2k_3) \quad \beta = (\varepsilon_{B_3}s^2 + s(\varepsilon_{B_1}k_3 + \varepsilon_{B_3}k_2) + \varepsilon_{B_1}k_2k_3) \quad (16)$$

As in the previous case, considering another set of parameters  $\mathbf{X}'$ , the following relationships are obtained:

$$\varepsilon_{B_1} = \varepsilon'_{B_1} \quad \varepsilon_{B_3} = \varepsilon'_{B_3} \quad \varepsilon_{B_1}k_2 + \varepsilon_{B_2}k_3 = \varepsilon'_{B_1}k'_2 + \varepsilon'_{B_2}k'_3 \quad k_2k_3 = k'_2k'_3$$

$$\varepsilon_{B_1}k_2k_3 = \varepsilon'_{B_1}k'_2k'_3 \quad \varepsilon_{B_1}k_3 + \varepsilon_{B_3}k_2 = \varepsilon'_{B_1}k'_3 + \varepsilon'_{B_3}k'_2 \quad k_2 + k_3 = k'_2 + k'_3 \quad (17)$$

Solving these equations, it is concluded that in all cases, the system is identifiable with a unique solution.

In order to check that the application of the ANN methodology leads us to these results in the case of the prediction of  $(k_2, k_3, \varepsilon_{B_1})$ ,  $(k_2, k_3, \varepsilon_{B_2})$  or  $(k_2, k_3, \varepsilon_{B_3})$ , 43  $A_T/t$  curves are generated from the values of the parameters obtained from the corresponding ED. The training of the ANN is performed with these curves, obtaining as optimal networks the one formed by 2 hidden layers with 1 neuron in the 1st layer and 2 in the 2nd layer (experience 1) or, 3 hidden layers with 4 neurons in each of the layers (experience 2).

Next, an  $A_T/t$  curve is generated for the following values:  $k_1 = 0.1 \text{ min}^{-1}$ ,  $k_2 = 0.05 \text{ min}^{-1}$ ,  $\varepsilon_{B_1} = 600 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_2} = 900 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_3} = 700 \text{ M}^{-1} \text{ cm}^{-1}$ . Subsequently, the different groups of parameters  $k_r$ ,  $\varepsilon_{B_j}$  are assumed unknown and the prediction is performed with the optimal ANN. The results obtained corroborate the conclusions obtained by the *deterministic analysis*, but in this mechanism, the difference between the values of the parameters with which the data had been generated and those calculated by ANN are significantly large. In this case, the values determined with ANN are used as initial estimates for the application of the mathematical optimization algorithm, which considerably improves the results (Table 3). The sequential application of both methods, ANN and mathematical optimization, constitutes the hybrid algorithm (ANN-MATOPT). Table 3 shows the obtained results using optimal ANNs in the prediction of  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_3})$  and the results obtained after subsequent mathematical optimization.

**Table 3** Obtained results in the prediction of  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_3})$  through ANN for the system  $B_2 \xrightarrow{k_2} B_1 \xrightarrow{k_3} B_1$  and in the subsequent mathematical optimization

Exp	ANN calculated values			Optimized values			SQD
	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.1203	0.0627	598.7500	0.1000	0.0500	600.000	4.8071E-31
2	0.0855	0.0436	598.1740	0.1000	0.0500	600.000	6.7540E-31
Exp	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.0832	0.0738	924.2574	0.1000	0.0500	900.000	5.4532E-31
2	0.1255	0.0537	898.1545	0.1000	0.0500	900.000	6.3530E-31
Exp	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\varepsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.0832	0.0738	724.2513	0.1000	0.0500	700.000	5.1232E-31
2	0.1255	0.0537	699.1053	0.1000	0.0500	700.000	5.3170E-31

#### 4.3.3 Model III $B_1 \xrightarrow{k_2} B_2 \xrightarrow{k_3} B_3$

In the case of this mechanism, the determination of the following parameters is studied:  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_3})$ . The Laplace transform of the total absorbance in the system is given by:

$$Y(s, \mathbf{X}) = \frac{(s \varepsilon_{B_1} + \varepsilon_{B_2} k_2 + \varepsilon_{B_3} k_3) [B_1]_0}{s^2 + s(k_2 + k_3)} \quad (18)$$

The following relationships between the parameters corresponding to two solutions  $\mathbf{X}$  and  $\mathbf{X}'$  are obtained from the *deterministic analysis*:

$$\varepsilon_{B_1} = \varepsilon'_{B_1} (\varepsilon_{B_2} k_2 + \varepsilon_{B_3} k_3) = \left( \varepsilon'_{B_2} k'_2 + \varepsilon'_{B_3} k'_3 \right) \quad k_2 + k_3 = k'_2 + k'_3 \quad (19)$$

Solving these equations, the subsequent conclusions are obtained: if the parameters to determine are  $(k_2, k_3, \varepsilon_{B_1})$  a unique solution is obtained and if  $(k_2, k_3, \varepsilon_{B_2})$  or  $(k_2, k_3, \varepsilon_{B_3})$  are determined, the system is non-identifiable since it has infinite solutions.

As in the previous mechanisms, in each of the cases to study the corresponding ED and the ANN training is carried out and the prediction process is performed in each of the cases. The optimal ANN for predicting the three sets of parameters are those consisting of 2 hidden layers with 6 neurons in the 1<sup>st</sup> layer and 7 neurons in the 2<sup>nd</sup> layer and 3 hidden layers with 4 neurons in each layer.

To carry out the prediction, an  $A_T/t$  curve is generated with the following parameter values:  $k_1 = 0.1 \text{ min}^{-1}$ ,  $k_2 = 0.05 \text{ min}^{-1}$ ,  $\varepsilon_{B_1} = 900 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_2} = 600 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_3} = 700 \text{ M}^{-1} \text{ cm}^{-1}$ . Subsequently, different groups of parameters are assumed unknown and they are determined through the application of the hybrid algorithm ANN-MATOPT. Table 4 shows the results obtained

**Table 4** Obtained results in the prediction of  $\mathbf{X} = (k_2, k_3, \epsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \epsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \epsilon_{B_3})$  through ANN for the system  $B_1 \xrightarrow{k_2} B_2 \xrightarrow{k_3} B_3$  and in the subsequent mathematical optimization

Exp	ANN calculated values			Optimized values			SQD
	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.0819	0.0615	891.6208	0.1000	0.0500	900.000	3.0810E-31
2	0.0974	0.0255	840.6093	0.1000	0.0500	900.000	6.2862E-31
3	0.0142	0.0631	939.2560	0.1000	0.0500	900.000	6.4095E-31
4	0.0950	0.1533	798.7227	0.1000	0.0500	900.000	6.5438E-31
5	0.0822	0.0565	888.6386	0.1000	0.0500	900.000	3.9470E-31
Exp	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.0145	0.1539	522.9659	0.1297	0.0293	770.4384	3.4576E-31
2	0.0738	0.1276	711.2020	-0.2045	0.3545	614.1030	2.5674E-31
3	0.0424	0.1277	488.3954	0.1274	0.0220	821.6612	2.4571E-31
4	0.0457	0.1297	547.9461	0.1320	0.0180	877.0900	4.1028E-31
5	0.1044	0.0278	435.2760	0.0195	0.1305	856.1416	3.5678E-31
6	0.0672	0.0223	452.8717	0.0351	0.1149	742.3933	3.2047E-31
7	0.1357	0.1509	452.8717	0.0125	0.1375	999.3013	4.6839E-31
8	0.0888	0.0749	481.6844	0.0181	0.1319	876.8066	2.7117E-31
Exp	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	$k_2/\text{min}^{-1}$	$k_3/\text{min}^{-1}$	$\epsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.0488	0.0345	663.2466	0.1122	0.0378	435.7853	5.1700E-31
2	0.0802	0.0210	706.8944	0.421	0.1079	607.3011	2.2187E-31
3	0.1302	0.0706	711.0206	1.1823	-1.0323	709.6872	5.3125E-29
4	0.0845	0.0502	724.1305	0.8255	-0.7755	712.8940	3.4836E-31
5	0.0435	0.0622	734.2410	0.5515	-0.4015	724.9080	8.2214E-30
6	0.1783	0.1171	725.7615	0.5375	-0.3875	725.8046	2.6378E-30
7	0.0846	0.0021	649.7490	0.1138	0.0362	424.0239	4.1908E-31
8	0.0345	0.1644	637.5440	0.0345	0.1166	614.2510	6.6560E-31

for the prediction of  $(k_2, k_3, \epsilon_{B_1})$ ,  $(k_2, k_3, \epsilon_{B_2})$  and  $(k_2, k_3, \epsilon_{B_3})$  for experiences performed with ANN of different architecture experiences and the results obtained after subsequent mathematical optimization.

The results obtained are in agreement with the deductions from the *deterministic analysis*, in the case of the prediction of  $(k_2, k_3, \epsilon_{B_1})$  in all of the experiences performed a unique solution is obtained and in the prediction of  $(k_2, k_3, \epsilon_{B_2})$  or  $(k_2, k_3, \epsilon_{B_3})$  it can be verified that the application of ANN and the subsequent optimization leads us to infinite solutions so the system is non-identifiable.

#### 4.3.4 Model IV $B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$

In the case of first-order consecutive irreversible reaction system, the prediction using ANN and subsequent optimization (hybrid algorithm ANN-MATOPT) of

the following set of parameters are studied:  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$ ,  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_2})$  and  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_3})$ .

The Laplace transform for the total absorbance and the corresponding relationships between the different set of parameters in this case are:

$$Y(s, \mathbf{X}) = \frac{(\varepsilon_{B_1} s^2 + s(\varepsilon_{B_1} k_2 + \varepsilon_{B_2} k_1) + \varepsilon_{B_3} k_1 k_2) [B_1]_0}{s^3 + s^2(k_1 + k_2) + sk_1 k_2} \quad (20)$$

$$\begin{aligned} \varepsilon_{B_1} &= \varepsilon'_{B_1} & k_1 + k_2 &= k'_1 + k'_2 & k_1 k_2 &= k'_1 k'_2 \\ \varepsilon_{B_1} k_2 + \varepsilon_{B_2} k_1 &= \varepsilon'_{B_1} k'_2 + \varepsilon'_{B_2} k'_1 & \varepsilon_{B_3} k_1 k_2 &= \varepsilon'_{B_3} k'_1 k'_2 \end{aligned} \quad (21)$$

Solving these equations for different situations allows us to conclude that if the set of parameters to determine are  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_1})$  or  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_3})$ , the system is identifiable and the solution is unique. In the case that the set of parameters to determine is  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_2})$  a three-equation system is obtained:

$$\varepsilon_{B_1} k_2 + \varepsilon_{B_2} k_1 = \varepsilon_{B_1} k'_2 + \varepsilon'_{B_2} k'_1 \quad k_1 + k_2 = k'_1 + k'_2 \quad k_1 k_2 = k'_1 k'_2 \quad (22)$$

Since the number of parameters to determine is three, the system is *identifiable*, but the solution is not unique. Two solutions are obtained,  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_2})$  and  $\mathbf{X}' = (k'_1, k'_2, \varepsilon'_{B_3})$  being the relationship between the two:

$$k'_1 = k_2 \quad k'_2 = k_1 \quad \varepsilon'_{B_2} = \varepsilon_{B_1} + \frac{k_1(\varepsilon_{B_2} - \varepsilon_{B_1})}{k_2} \quad (23)$$

In the same way as in the previous cases to study, the corresponding ED is performed, the ANN training is carried out and the prediction process is realized with the ANN that provide better results.

In the case that the parameters to determine are  $(k_1, k_2, \varepsilon_{B_1})$  the most appropriate ANN for the prediction are: 1 hidden layer with 10 neurons, 1 hidden layer with 15 neurons, 2 hidden layers with 15 neurons in each layer and 2 hidden layers with 20 neurons in each layer (experiences 1, 2, 3 and 4 respectively). If the prediction of  $(k_1, k_2, \varepsilon_{B_3})$  is performed, the most appropriate ANN for the prediction are: 1 hidden layer with 15 neurons, 2 hidden layers with 5 neurons in each one, 2 hidden layers with 7 neurons in the first layer, 5 in the second one and 2 hidden layers with 20 neurons in each layer (experiences 1, 2, 3 and 4 respectively).

In both cases, in order to carry out the prediction, an  $A_T/t$  curve is generated with the following values  $k_1=0.1 \text{ min}^{-1}$ ,  $k_2=0.05 \text{ min}^{-1}$ ,  $\varepsilon_{B_1}=900 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_2} \text{Lo}=600 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\varepsilon_{B_3}=700 \text{ M}^{-1} \text{ cm}^{-1}$ . Then,  $(k_1, k_2, \varepsilon_{B_1})$  or  $(k_1, k_2, \varepsilon_{B_3})$  are assumed unknown and through the application of the different ANN the prediction of the parameters is performed. Subsequently, the mathematical optimization is applied to improve the results obtained through ANN, the results obtained lead us to the same conclusions as the deterministic analysis; when determining the

**Table 5** Obtained results in the prediction of  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_1})$  and  $\mathbf{X} = (k_1, k_2, \varepsilon_{B_3})$  through ANN for the system  $B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$  and in the subsequent mathematical optimization

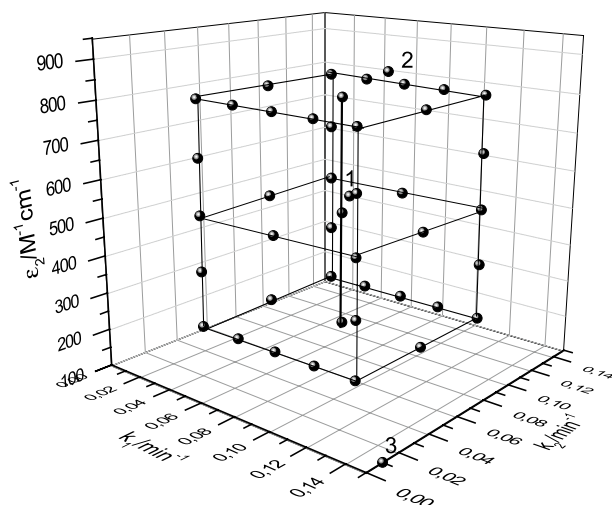
Exp	ANN calculated values			Optimized values			SQD
	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.08251	0.07173	895.3802	0.10000	0.05000	900.0000	4.4373E-31
2	0.09105	0.06087	893.4207	0.10000	0.05000	900.0000	2.5884E-31
3	0.06647	0.07347	899.3159	0.10000	0.05000	900.0000	4.8071E-31
4	0.08513	0.06342	893.8493	0.10000	0.05000	900.0000	4.0675E-31
Exp	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.07619	0.07556	711.0126	0.10000	0.05000	700.0000	4.5606E-31
2	0.07101	0.07925	799.9181	0.10000	0.05000	700.0000	4.5606E-31
3	0.11184	0.09352	699.0455	0.10000	0.05000	699.9999	5.5467E-31
4	0.09664	0.02637	731.4196	0.10000	0.05000	699.9999	5.4234E-31

parameters  $(k_1, k_2, \varepsilon_{B_1})$  or  $(k_1, k_2, \varepsilon_{B_3})$  the system is identifiable, obtaining a unique solution (Table 5).

Lastly, we have studied the joint determination of  $(k_1, k_2, \varepsilon_{B_2})$ , according to the *deterministic analysis* when these parameters are jointly determined from total absorbance kinetic data, the system is identifiable but the solution is not unique, obtaining two solutions related to each other by the Eq. (17). Following the same procedures as in the previous cases, the corresponding ED is carried out and a series of  $A_T/t$  curves are generated with which the training that provides the optimal networks is performed, a network formed by 2 hidden layers with 4 neurons in each layer. With the optimal ANN, considering  $\varepsilon_{B_1}$  and  $\varepsilon_{B_3}$  as known values, we carry out the prediction of  $k_1, k_2$  and  $\varepsilon_{B_2}$ .

Considering the point 1 of Fig. 6 that is found inside the 3D figure of the corresponding ED and which corresponds to the values  $k_1=0.1 \text{ min}^{-1}$ ,  $k_2=0.05 \text{ min}^{-1}$ ,  $\varepsilon_{B_2}=600 \text{ M}^{-1} \text{cm}^{-1}$ , a curve  $A_T/t$  is generated with these values and considering that  $\varepsilon_{B_1}=900 \text{ M}^{-1} \text{cm}^{-1}$  and  $\varepsilon_{B_3}=700 \text{ M}^{-1} \text{cm}^{-1}$ . If  $k_1, k_2$  and  $\varepsilon_{B_2}$  are unknown, according to the *deterministic analysis*, there exists two sets of parameters that can fit the data  $A_T/t$ :  $\mathbf{X}=(k_1=0.1 \text{ min}^{-1}, k_2=0.05 \text{ min}^{-1}, \varepsilon_{B_2}=600 \text{ M}^{-1} \text{cm}^{-1})$  and  $\mathbf{X}'=(k'_1=0.05 \text{ min}^{-1}, k'_2=0.1 \text{ min}^{-1}, \varepsilon'_{B_2}=300 \text{ M}^{-1} \text{cm}^{-1})$ . The results obtained in the prediction of these parameters through the application of ANN and the subsequent mathematical optimization are shown in Table 6. The values provided by ANN are far from the real values, so we use the values as initial estimates for the mathematical optimization process. The results obtained through the application of the hybrid algorithm ANN-MATOPT show that two solutions that are mathematically possible are obtained. The values obtained for  $\varepsilon_{B_2}$  by ANN is the one that conditions that the mathematical optimization process leads us to one or another solution, the two solutions are mathematically acceptable, performing the statistical analysis of residuals there are no appreciable differences between the two.

Points 2 and 3 of Fig. 6 are outside the ED and correspond respectively to the parameter values  $(k_1=0.15 \text{ min}^{-1}, k_2=0.01 \text{ min}^{-1}, \varepsilon_{B_2}=100 \text{ M}^{-1} \text{cm}^{-1})$  and



**Fig. 6** Three-dimensional experimental design, 3D (3 factors  $k_1$ ,  $k_2$  and  $\varepsilon_2$ ) for the system  $B_1 \rightarrow B_2 \rightarrow B_3$ . Points 1, 2 and 3 indicate the experiences in which the values of  $k_1$ ,  $k_2$  and  $\varepsilon_2$  are unknown and will be determined

**Table 6** Obtained results in the prediction of  $\mathbf{X} = (k_2, k_3, \varepsilon_{B_1})$  through ANN for the system  $B_1 \rightarrow^{k_1} B_2 \rightarrow^{k_2} B_3$  and in the subsequent mathematical optimization (point 1)

Exp	ANN calculated values			Optimized values			SQD
	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\varepsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.0420	0.0718	456.0560	0.0500	0.1000	300.0000	7.5188E-31
2	0.0800	0.1618	323.8090	0.0500	0.1000	300.0000	5.7931E-31
3	0.0348	0.0640	482.6900	0.0500	0.1000	300.0000	6.9025E-31
4	0.0717	0.1347	435.3400	0.0500	0.1000	300.0000	6.1458E-31
5	0.0084	0.0359	498.2990	0.1000	0.5000	600.0000	6.9025E-31
6	0.0431	0.0372	512.9900	0.1000	0.5000	600.0000	4.1908E-31
7	0.1023	0.0299	614.2844	0.1000	0.5000	600.0000	7.1419E-31
8	0.1370	0.1083	508.4349	0.1000	0.5000	600.0000	5.9160E-31

( $k_1=0.15 \text{ min}^{-1}$ ,  $k_2=0.01 \text{ min}^{-1}$ ,  $\varepsilon_{B_2}=950 \text{ M}^{-1} \text{cm}^{-1}$ ), in the case of the jointly determination of  $k_2$ ,  $k_3$  and  $\varepsilon_{B_2}$  from the  $A_T/t$  data, in each case there is another possible set of parameters that represent the system ( $k'_1=0.01 \text{ min}^{-1}$ ,  $k'_2=0.15 \text{ min}^{-1}$ ,  $\varepsilon'_{B_2}=11,000 \text{ M}^{-1} \text{cm}^{-1}$ ) and ( $k'_1=0.01 \text{ min}^{-1}$ ,  $k'_2=0.15 \text{ min}^{-1}$ ,  $\varepsilon'_{B_2}=1650 \text{ M}^{-1} \text{cm}^{-1}$ ).

Table 7 collects the most significant results obtained after the application of the hybrid algorithm in these cases, which confirms the applicability of the ANN-MATOPT algorithm for the determination of  $k_1$ ,  $k_2$  and  $\varepsilon_{B_2}$ , although the values of these parameters are outside of the ED, it is also checked that the two solutions

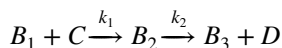
**Table 7** Obtained results in the prediction of  $\mathbf{X} = (k_2, k_3, \epsilon_{B_3})$  through ANN for the system  $B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$  and in the subsequent mathematical optimization (points 2 and 3)

Exp	ANN calculated values			Optimized values			SQD
	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.1347	0.1317	791.2403	0.1500	0.0100	950.0000	8.8747E-31
2	0.0165	0.0922	818.1860	0.0100	0.1500	1650.0000	6.0397E-31
1	0.0388	0.0479	273.4312	0.1500	0.0100	100.0000	4.4373E-31
2	0.0234	0.1765	134.5648	0.01	0.1500	- 111,000.0000	5.6608E-31

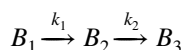
obtained are acceptable from a mathematical point of view although not from a chemical point of view ( $\epsilon'_{B_2} < 0$ ).

Finally, the proposed methodology (ANN-MATOPT) has been used to process experimental kinetic data from the reaction between carbonyl cyanide 3-chlorophenylhydrazone ( $B_1$ ) with 2-mercaptoethanol (C), whose kinetics has been studied in different works [28–30].

The mechanism of this reaction consists of two consecutive irreversible reactions, in the first step, an intermediate adduct ( $B_2$ ) is formed, which is then hydrolyzed in an intramolecular reaction to give the product 3-chlorophenylhydrazonocyanacetamide, ( $B_3$ ) and the subproduct ethylene sulphide (D).



The reaction is carried out in an excess of the reagent C, so a pseudo first order can be assumed and it is possible to express it schematically according to the IUPAC norms [19]:



Chau et al. [29] have performed a classic kinetic study at temperature of 20.0° C and pH = 4.3, following the evolution of the reaction by measuring the total absorbance (AT) at different wavelengths. The values obtained by these authors for the rate constants are:  $k_1 = 0.33600 \text{ min}^{-1}$  and  $k_2 = 0.02004 \text{ min}^{-1}$ . We use the experimental data obtained when the reaction is followed at  $\lambda = 375 \text{ nm}$  to determine by the hybrid algorithm ANN-MATOPT the kinetic constants and the molar absorption coefficients at that  $\lambda$ . The values of the molar absorption coefficients of each of the species at  $\lambda = 375 \text{ nm}$  are:  $\epsilon_{B_1} = 1.63 \times 10^4 \text{ mol}^{-1} \text{ cm}^{-1} \text{L}$ ,  $\epsilon_{B_2} = 2.00 \times 10^4 \text{ mol}^{-1} \text{ cm}^{-1} \text{L}$  and  $\epsilon_{B_3} = 1.48 \times 10^4 \text{ mol}^{-1} \text{ cm}^{-1} \text{L}$ .

Based on the values found in the bibliography, the interval variation of the parameters is established to perform the ED:  $[0.35\text{--}0.02 \text{ min}^{-1}]$  for the constants ( $k_i$ ) and  $[3 \times 10^4\text{--}1 \times 10^4 \text{ Lmol}^{-1} \text{ cm}^{-1}]$  for the molar absorption coefficients ( $\epsilon_j$ ). Once the corresponding ED has been performed, 51 kinetic curves  $A_T/t$  were generated with which the ANN training has been carried out and the optimal network has been obtained with which the prediction of the following parameters has been carried out:



**Table 8** Obtained results in the prediction of  $\mathbf{X} = (k_1, k_2, \epsilon_{B_1})$ ,  $\mathbf{X} = (k_1, k_2, \epsilon_{B_1})$  and  $\mathbf{X} = (k_1, k_2, \epsilon_{B_2})$  for the system  $B_1 \xrightarrow{k_1} B_2 \xrightarrow{k_2} B_3$  and in the subsequent mathematical optimization (experimental data)

Exp	ANN calculated values			Optimized values			SQD
	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_1}/\text{M}^{-1} \text{cm}^{-1}$	
1	0.3009	0.3207	1.6305E04	0.3360	0.0200	1.63E04	1.3540E-08
2	0.1693	0.2010	1.6360E04	0.3360	0.0200	1.63E04	5.3460E-08
3	0.1854	0.1782	1.6992E04	0.3360	0.0200	1.63E04	1.3780E-08
4	0.3035	0.1529	1.6345E04	0.3360	0.0200	1.63E04	5.6543E-08
Exp	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_3}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.1993	0.4217	1.4847E04	0.3363	0.0200	1.4795E04	2.9376E-08
2	0.2850	0.1988	1.4601E04	0.3363	0.0200	1.4795E04	2.9376E-08
3	0.2569	0.3824	1.4826E04	0.3363	0.0200	1.4795E04	3.5728E-08
4	0.1749	0.4234	1.4699E04	0.3363	0.0200	1.4795E04	5.5654E-08
Exp	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	$k_1/\text{min}^{-1}$	$k_2/\text{min}^{-1}$	$\epsilon_{B_2}/\text{M}^{-1} \text{cm}^{-1}$	SQD
1	0.2523	0.4836	2.4555E04	0.3362	0.0200	2.0001E04	3.1679E-08
2	<i>0.0424</i>	<i>0.6062</i>	<i>2.6167E04</i>	<i>0.0200</i>	<i>0.3362</i>	<i>7.8406E04</i>	3.1679E-08
3	0.1449	0.0178	2.6858E04	0.3362	0.0200	2.0001E04	3.1679E-08
4	<i>0.0896</i>	<i>0.3938</i>	<i>2.3711E04</i>	<i>0.0200</i>	<i>0.3362</i>	<i>7.8406E04</i>	3.1679E-08

$(k_1, k_2, \epsilon_{B_1})$ ,  $(k_1, k_2, \epsilon_{B_2})$  and  $(k_1, k_2, \epsilon_{B_3})$ . Table 8 shows the results obtained in different experiences after the application of the hybrid algorithm ANN-MATOPT for the prediction of the different set of parameters. As it occurs in the treatment of synthetic data, when  $(k_1, k_2, \epsilon_{B_1})$  or  $(k_1, k_2, \epsilon_{B_3})$  are determined the system is identifiable with a unique solution and when the parameters to determine are  $k_1$ ,  $k_2$  and  $\epsilon_{B_1}$  the system is identifiable, but the solution is not unique, obtaining two solutions that are mathematically acceptable. Table 8 shows the two solutions, the real one and the other one that is mathematically acceptable (indicated in italics).

## 5 Conclusion

In this work the hybrid algorithm ANN-MATOPT has been used to determine in different reaction mechanisms, the rate constants ( $k_i$ ) of the different stages and the molar absorption coefficients ( $\epsilon_j$ ) of the different species participating in the reaction.

In a first stage the ANN methodology is applied, the analysis of the parameters given by ANN (regressions *outputs/targets*, MSE, etc,..) show that in some of the cases, the obtained results are unacceptable, telling the need to apply, in the second step, a mathematical optimization method to improve the results obtained with ANN and that originate the hybrid algorithm ANN-MATOPT.

The advantage of applying the ANN methodology over other parameter determination methods, is that it is a treatment that does not need to start from initial

estimates, so it is not necessary to know the magnitude order of the parameters to determine. The mathematical optimization methods need estimated values of the parameters and it is also convenient that these estimations are close to the real values for the optimization to be carried out correctly. In the proposed methodology, ANN is applied initially, so that the results obtained can be used as initial estimates of the second method (mathematical optimization) that does require the use of previous values for the parameters to determine. It is observed that if the mechanism is simple (model I) the result provided by ANN are acceptable and in no case is necessary to apply subsequently the mathematical optimization method. For more complex mechanisms (II, III and IV) in all situations, it is necessary the application of the mathematical optimization so that the results can be considered acceptable.

When parameters of distinct nature are determined from kinetic data from the measurement of a total property, problems of ambiguity on the prediction may arise, that is, there is more than one group of parameters that represents the experimental data. The application of the Laplace transform has allowed us to detect beforehand the existence of ambiguity for the studied mechanism. The application of the ANN methodology in the case of situations where there is ambiguity or equivalence in the parameters, leads us to the same conclusions as the *deterministic analysis* and the mathematical optimization methods. ANN leads us to any of the acceptable solution from the mathematical point of view. In the cases of mechanisms I and II, a group of parameters is obtained that fits the kinetic data in all cases, while for mechanism III, it depends on the parameters that are determined. If  $(k_2, k_3, \epsilon_{B_1})$  are determined, a group of parameters is obtained and for  $(k_2, k_3, \epsilon_{B_2})$  and  $(k_2, k_3, \epsilon_{B_3})$ , an infinite number of parameters that fit the kinetic data, non-identifiable system is obtained.

In the mechanism IV, it happens the same as the previous case, in the determination of  $(k_1, k_2, \epsilon_{B_1})$  and  $(k_1, k_2, \epsilon_{B_2})$  a unique solution is obtained, while when determining  $(k_1, k_2, \epsilon_{B_3})$  we find two indiscernible solutions from each other, equally valid from a mathematical point of view. It is worth to point out the influence of the molar absorption coefficient of the intermediate species obtained with ANN since its value has a decisive influence on the hybrid algorithm ANN-MATOPT leading us to one or the other solution.

The application of deterministic analysis in the mechanisms studied indicates that in the special case  $k_1 = k_j$  there is no "slow-fast" ambiguity in the determination of parameters.

In all cases in which there is ambiguity, that is, in those systems identifiable with a finite number of solutions or non-identifiable with infinite solutions, all of them are acceptable from a mathematical point of view, but not so from a chemical point of view. All the sets of parameters fit the experimental data, but only one of them contains the true parameters. In these cases, all the solutions are mathematically acceptable, performing a statistical analysis of residuals, there are no appreciable differences found between them, in order to discern which is the true solution we must resort to other considerations different from the mathematical or statistical, such as those of the strictly chemical type. The molar absorption coefficient of a certain specie at a wavelength  $\lambda(\epsilon_j)$ , is a parameter that has a characteristic value and can be used for the chemical identification of that specie. Similarly, the value of the rate constants of the different stages of a reaction mechanism ( $k_i$ ), have a determined

value under certain reaction conditions (temperature, pH, ...). For this reason, only a certain set of  $k_i$  and  $\epsilon_j$  values are chemically true, although there is more than one set of parameters that fit the time total absorbance data and leads to a minimum value of the sum of the quadratic deviations (SQD).

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. F. Pérez Pla, J.F. Bea Redón, R. Valero, *Chemom. Intell. Lab. Syst.* **53**, 1 (2000)
2. B. Svir, O.V. Klymenko, M.S. Platz, *Comput. Chem.* **26**, 379 (2002)
3. M.M. Canedo, J.L. González-Hernández, *Chemom. Intell. Lab. Syst.* **66**, 63 (2003)
4. M.M. Canedo, J.L. González-Hernández, *J. Math. Chem.* **49**, 163 (2011)
5. M.M. Canedo, J.L. González-Hernández, *S. Encinar del Dedo, App. Math. and Comp.* **219**, 7089 (2013)
6. E. Bezemer, S.C. Rutan, *Chemom. Intell. Lab. Syst.* **59**, 19 (2001)
7. S. Bijlsma, H. Boelens, H. Hoefsloot, A.K. Smilde, *Anal. Chim. Acta* **419**, 197 (2000)
8. B. Kovacs, J. Tóth, *Int. J. Appl. Math. Comput. Sci.* **4**, 7 (2007)
9. N.H.T. Lemes, E. Borges, J.P. Braga, *Chemom. Intell. Lab. Syst.* **96**, 84 (2009)
10. F. Amato, J.L. González-Hernández, *J. Havel, Talanta* **93**, 72 (2012)
11. M.M. Canedo, J.L. González-Hernández, *S. Encinar del Dedo, J. Math. Chem.* **51**, 1634 (2013)
12. J.L. González-Hernández, M.M. Canedo Alonso, *S. Encinar del Dedo, MATCH Commun. Math. Comput. Chem.* **79**, 619 (2018)
13. S. Encinar del Dedo, J.L. González-Hernández, M.M. Canedo, *MATCH Commun. Math. Comput. Chem.* **72**, 427 (2014)
14. S. Encinar del Dedo, J.L. González-Hernández, M.M. Canedo, D. Juanes, *J. Math. Chem.* **53**, 1080 (2015)
15. J.L. González-Hernández, M.M. Canedo Alonso, *S. Encinar del Dedo, MATCH Commun. Comput. Chem.* **83**, 295 (2020)
16. S. Vadjá, H. Rabitz, *J. Phys. Chem.* **98**, 5265 (1994)
17. A. Balogh, G. Lente, J. Kalmár, I. Fábíán, *Int. J. Chem. Kinet.* **47**, 773 (2015)
18. A.I. Petrov, V.D. Dergachev, *Int. J. Chem. Kinet.* **49**, 494 (2017)
19. K.J. Laidler, *Pure Appl. Chem.* **68**, 149 (1996)
20. A. Tarek, E.A. Hopgood, L. Nolle, A. Battersby, *Eng. Lett.* **13**(2), 124 (2006)
21. A. Freeman, D.M. Skapura, *Neural Networks, Algorithms, Applications and Programming Techniques*, 1st edn. (Adinson-Wesley, Massachusetts, 1991)
22. J. Zupan, J. Gasteiger, *Neural Networks for Chemists. An Introduction*, 1st edn. (VCH Weinheim, New York, 1993)
23. J.R. Hilera, V.J. Martínez, *Redes Neuronales Artificiales: Fundamentos, Modelos y Aplicaciones*, 1st edn. (Alfaomega, Madrid, 2000)
24. P.R. Abdy, M.A.H. Dempster, *Introduction to Optimization Methods*, 1st edn. (Chapman and Hall, Cambridge, 1974)

25. M.A. Wolfe, *Numerical Methods for Unconstrained Optimization: An Introduction*, 1st edn. (Van Nostrand Reinhold, New York, 1978)
26. MATLAB & Simulink, © 1994–2021, The MathWorks, Inc
27. P. Gemperline, *Practical Guide to Chemometrics*, 2nd edn. (CRC Press, Boca Raton, 2006)
28. R.H. Bisby, E.W.K. Thomas, *J. Chem. Ed.* **63**(11), 990 (1986)
29. F.T. Chau, K.W. Mok, *Comput. Chem.* **16**, 239 (1992)
30. S. Bijlsma, D.J. Louwerse, A.K. Smilde, *J. Chemom.* **13**, 311 (1999)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.