Algorithm and software for modelling of food protein hydrolysis kinetic

Abakarov, A.[a]

[a]*Universidad Politécnica de Madrid, Madrid, Spain (alik.abakarov@usm.cl)*

ABSTRACT

The objective of this research was to propose two numerical approaches and based on these approaches user-friendly dialogue software for determining the kinetics of enzyme hydrolysis of protein at different initial reaction parameters, which could be an initial substrate and enzyme concentrations, temperature and pH-value. The proposed approaches mainly are based on the utilization of adaptive random search technique, Artificial Neural Networks (ANNs) and the following well-known *exponential kinetic equation* presented a relationship between reaction rate and enzymatic reaction parameters $dh/dt = aexp(-bh)$, where coefficients a and b have different expressions according to different reaction mechanism. The first approach consists of using the adaptive random search technique for fitting the kinetic constants of enzymatic reactions to a set of given experimental or time course profiles (TCPs) in the form of degree of hydrolysis over reaction time. Use of this technique can give much greater confidence that if a global optimal solution had existed within the global domain then it would not have been missed by the search routine. The second approach uses ANNs for estimating the values of coefficients a and b of the exponential kinetic equation. In this case each of experimental TCPs is fitted to this equation by using adaptive random search technique. Optimal a and b values obtained for TCPs and initial enzymatic reaction parameters are used as an input data for the ANNs learning process. The user-friendly dialogue software "ANNEKs" (**ANN E**nzyme **K**inetic**s**) realized two mentioned above approaches is developed. Findings from the work reported in this study would suggest that the developed user-friendly interfaces and utilized numerical approaches make the "ANNEKs" software package useful for food scientists and engineers.

Keywords: protein hydrolysis kinetic; exponential kinetic equation; adaptive random search; Artificial Neural Network; sophisticated software.

INTRODUCTION

Enzymatic hydrolysis of food proteins is well-known method to modify and improve their functional and nutritional properties [1, 2]. It provides desirable characteristics from the food technology point of view. Protein hydrolysates are typically employed as food ingredients to stabilize interfaces, bind water, lipids and aroma, solubilise other ingredients or improve organoleptic features of the product such as colour, odour and flavor [1, 2]. It is known that the protein structure is sensitive to processing conditions [1]. Therefore, in approaching the process design for the enzyme hydrolysis of proteins, it is important to be able to predict processing outputs in response to different input variables, such as substrate and enzyme initial concentrations, and hydrolysis time. A useful tool for this purpose is an accurate mathematical model, however, in many cases it is difficult to construct a proper model for an enzymatic reaction since the constructed kinetic model obeys an enzyme with reaction conditions in which the kinetic constants were estimated only. At the same time, the determination of kinetic constants for other required reaction conditions could be extremely difficult due to their time consuming procedures and cost. An alternative approach to obtaining a model of the reaction kinetics would be the use of Artificial Neural Networks (ANNs), since it is well known that ANNs can work as universal approximator of non-linear functions and, basing on the experimental data, can be successfully used in assessing the dynamics of complex processes.

It is known, that well designed dialogue software significantly simplifies an engineering calculations process from its initial stage consisted of problem definition or (and) parameterization to the final stage consisted of realizing required computations. Thus, the objective of this research consists of development algorithms and graphic user interface (GUI) software package to meet such food engineering need as determining the kinetics of enzyme hydrolysis of protein at different initial reaction parameters.

MATERIALS & METHODS

Exponential kinetic equation

The following well-known exponential kinetic equation presents a relationship between the reaction rate *R*, and the concentrations of initial substrate S_0 and enzyme E_0 [3]:

$$\frac{dDH}{dt} = a\exp(-bDH) \tag{1}$$

where the coefficients *a* and *b* have different expressions according to different reaction mechanisms. For example, the following expressions of coefficients *a* and *b* can be used for the mechanism of substrate-inhibition [3]:

$$a = \frac{k_2 k_s E_0}{S_0 k_s + S_0^2}\ ,\ b = \frac{k_3 k_m k_s}{k_2(k_s + S_0)}\ ,\ k_3 = \frac{k_m}{k_d}$$

where: k_m is M-M constant,
k_s is a substrate inhibition constant,
k_2 is a reaction rate constant of enzymatic hydrolysis,
k_d is a inactivation constant of enzyme.

Determination of the kinetics of enzyme hydrolysis

The two following approaches are used in this study for determining the kinetics of enzyme hydrolysis of protein at different initial reaction parameters:

1. This approach consists of estimating the parameters of exponential kinetic equation. The best-fitting parameters (for example, kinetic constants k_m, k_s, k_2 , k_d) are estimated from time course profiles as a regression problem. Namely, an adaptive random search algorithm is used to minimize the sum of squared residuals. Use of this method can give much greater confidence than if a true global optimal solution had existed within the global domain (better than local optima), since then it would not have been missed by the search routine.
2. This approach consists of using the ANNs for estimating the coefficients *a* and *b* of exponential kinetic equation. In this approach each of the experimental TCPs are fitted to exponential equation by an adaptive random search technique. Optimal *a* and *b* values obtained for TCPs, and different initial reaction parameters (for example, initial enzyme and substrate concentrations) are used as a given and (input) data for the ANNs learning process.

Adaptive random search algorithm

The adaptive random search algorithm belongs to a specific class of global stochastic optimization algorithms [6]. This class of algorithms is based on generating the decision variables from a given probability distribution, and the term ‘‘adaptive” consists of modifications to the probability distribution utilized in the searching process, which, throughout the whole search process, act as minimum computations of the objective function, locating global solutions. The pedestal probability distribution is utilized in the adaptive random search. After every calculation of objective function, the pedestal distribution of decision variables is modified so that the probability of finding the optimal value of the objective function is increased. For example, Fig. 1 shows a pedestal frequency distribution for the two-dimensional case of an optimization problem can be obtained in the middle of the search process.

“ANNEKs” software

Borland C++ Builder 6.0 was used to design the “ANNEKs” software. “ANNEKs” contains three worksheets oriented to: a) estimating the kinetic parameters of exponential kinetic equation (Fig. 2), b) using the ANNs for estimating the coefficients *a* and *b* of exponential kinetic equation (Fig. 3). The worksheet “Simultaneous optimization” (Fig. 2) presents the following functionality: definition of number, bounds and initial values of variables used in the exponential kinetic equation (1); use the mathematical expression parser to define expressions for coefficients *a* and *b* according to the required reaction mechanisms; input of the experimental TCPs from text files and its graphical representation; computation of the MSE-values and Pearson correlation coefficients for each of the TCPs; adaptive random search algorithm parametrization.

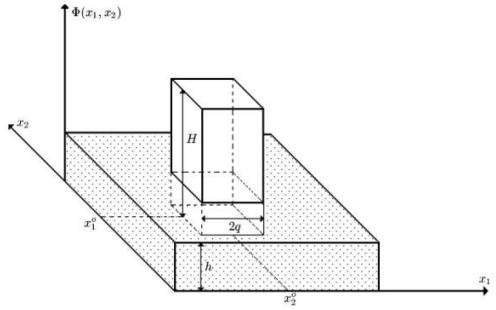

Figure 1. Pedestal frequency distribution for a two-dimension case.

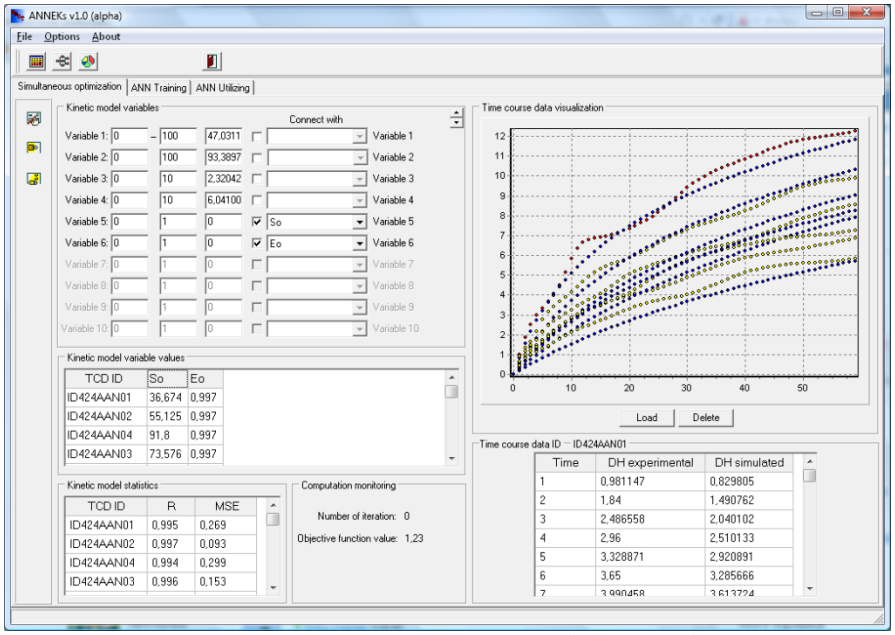

Figure 2. Simultaneous optimization worksheet of ANNEKs software.

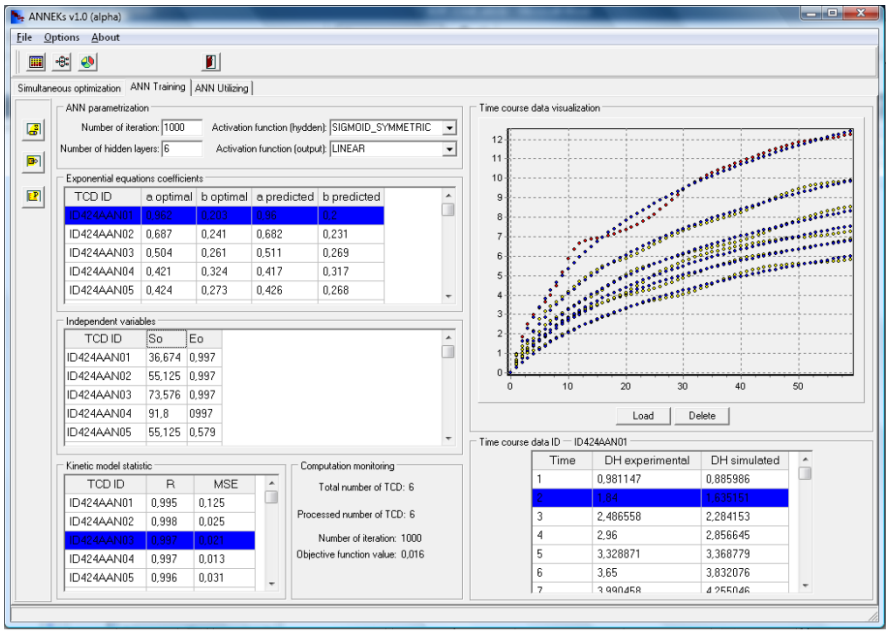

Figure 3. Worksheets “ANN Training” and “ANN Utilizing” oriented to estimation of a and b coefficient by using ANNs.

The worksheets “ANN Training” and “ANN Utilizing” (Fig. 3) present the following functionality: input of the experimental TCPs from text files and its graphical representation; computation of optimal values of coefficients *a* and *b* for each of the TCDs; ANNs parametrization consisted of choosing a number of neurons in hidden layer, maximum number of epochs, activation functions type for the hidden and output layers; definition of values of initial reaction parameters; ANN’s input data set definition; computation of the MSE-values and Pearson correlation coefficients for each of the TCPs.

RESULTS & DISCUSSION

Six experimental time course profiles in the form of degree of hydrolysis over reaction time are shown in Figure 4 [4, 5, 7].

First approach

The results from performance tests on models based on equations (1) and (2) are shown in Figure 2. In these figures, TCPs predicted by an exponential kinetic model (1) and (2) are compared with the experimental TCPs obtained from the six enzyme hydrolysis experiments. The optimal values found by “ANNEKs” software are as follows: $k_m = 2.32$, $k_s = 93.39$, $k_2 = 47.03$, $k_3 = 6.04$. Figure 2 shows lack of agreement between the predicted outcomes and experimental data. This lack of fit outside the range from which experimental data were taken is a clear indication that the exponential kinetic model (1) and (2) does not adequately describe all the complex reactions involved. The kinetic constants k_m, k_s, k_2, k_3 are clearly not constants, but vary as the reaction progresses. It might be possible to resolve this complexity by using some functional relationship that would express each of the kinetic constants as a function of initial substrate and enzyme concentration. But, in this case we are still confronted with the problem of determining such a relationship, which we do not know because of the complexity of the reaction mechanism.

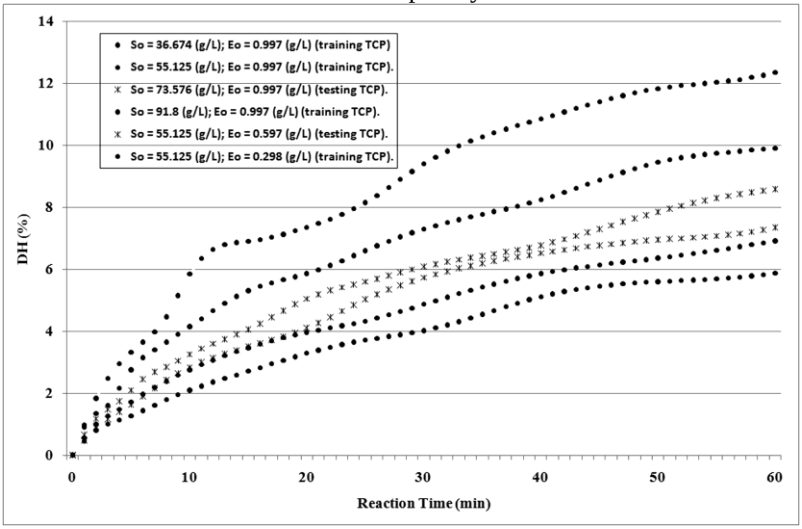

Figure 4. Six experimental TCPs.

Second approach

The experimental TCPs were divided into two groups. The first group consisted of TCPs for use in the training, and the second group consisted of TCPs used for testing model performance (Fig. 4). The optimal values of coefficients *a* and *b* computed by “ANNEKs” for TCPs of both groups are presented in Table 1.

Table 1. Optimal values of coefficients *a* and *b* computed by “ANNEKs” software

	Initial substrate and enzyme concentrations (g/L)					
	TCPs of learning group				**TCPs of testing group**	
	So = 36.674	**So = 55.125**	**So = 55.125**	**So = 91.8**	**So = 73.576**	**So = 55.125**
Coefficients	**Eo = 0.997**	**Eo = 0.997**	**Eo = 0.298**	**Eo = 0.997**	**Eo = 0.997**	**Eo = 0.597**
a	0.93995	0.65853	0.28486	0.43759	0.50950	0.4251
b	0.19774	0.23535	0.30837	0.33139	0.2640	0.2735

Figure 5 depicts the TCPs of both groups: training and testing along with the smooth curves corresponding to the optimal values of coefficients *a* and *b* from Table 1.

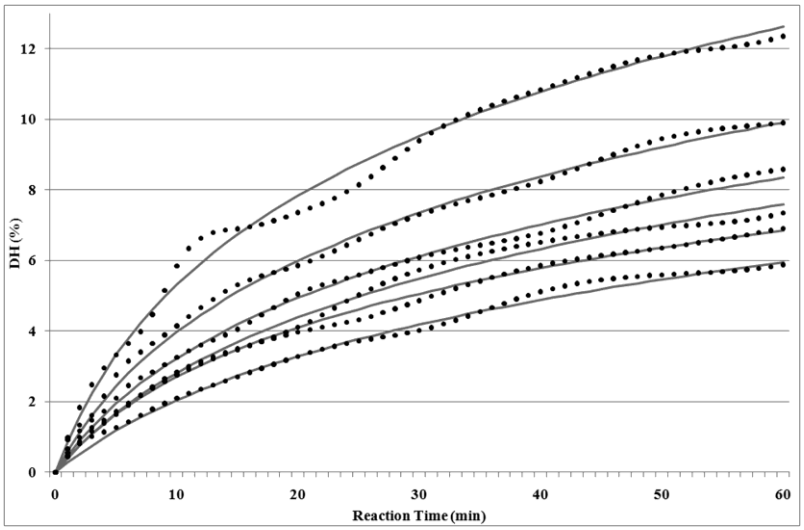

Figure 5. TCPs along with the smooth curves corresponding to the optimal values of coefficients *a* and *b*.

The values of coefficients *a* and *b* obtained from ANN model constructed by using ANNEKs software are shown in Table 2.

Table 2. Values of coefficients *a* and *b* obtained by the ANN model.

	Initial substrate and enzyme concentrations (g/L)					
	TCPs of learning group				**TCPs of testing group**	
	So = 36.674	**So = 55.125**	**So = 55.125**	**So = 91.8**	**So = 73.576**	**So = 55.125**
Coefficients	**Eo = 0.997**	**Eo = 0.997**	**Eo = 0.298**	**Eo = 0.997**	**Eo = 0.997**	**Eo = 0.597**
a	0.93977	0.658826	0.28475	0.43756	0.5082	0.4354
b	0.19769	0.23538	0.30830	0.33141	0.2703	0.2796

The TCPs used for testing are shown in Figure 6, along with the two smooth curves corresponding to the *a* and *b* values predicted by the ANN model. Table 3 summarizes information about the correlation coefficients and MSE values between the experimental TCPs and the curves obtained from the trained ANN-based model.

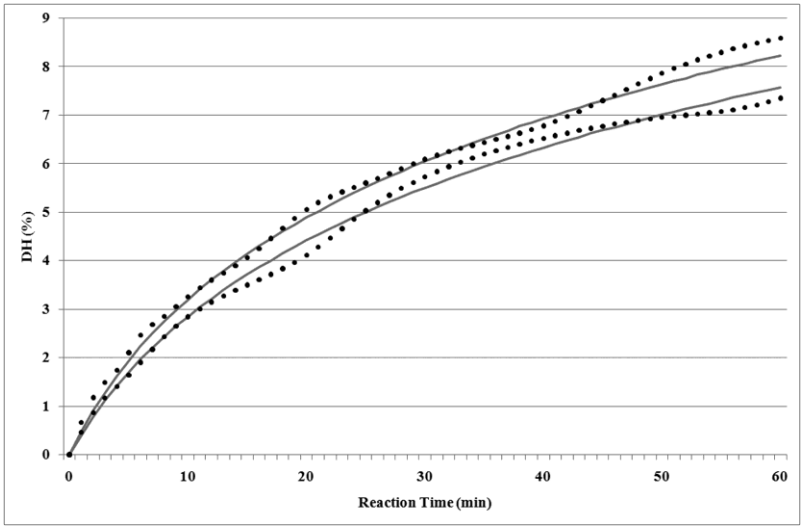

Figure 6. TCPs used for testing along with the smooth curves predicted by the ANN model.

The results obtained from ANN approach indicate its high predictive performance for processes involving complex reaction kinetics [7].

Table 3. Correlation coefficients and MSE values between experimental the TCPs and the predicted curves

	Initial substrate and enzyme concentrations (g/L)					
	TCPs of learning group				**TCPs of testing group**	
	So = 36.674	**So = 55.125**	**So = 55.125**	**So = 91.8**	**So = 73.576**	**So = 55.125**
	Eo = 0.997	**Eo = 0.997**	**Eo = 0.298**	**Eo = 0.997**	**Eo = 0.997**	**Eo = 0.597**
R^2	0.996	0.998	0.997	0.997	0.997	0.996
MSE	0.0938	0.0267	0.0185	0.0152	0.0309	0.0329

CONCLUSION

Findings from the work reported in this study would suggest the following conclusions:

- The first approach has limited applicability to the enzyme hydrolysis of protein, because they may show lack of agreement between the predicted outcomes and experimental data.
- The second approach can be very effective in developing predictive models for processes involving complex reaction kinetics that would otherwise be difficult to develop by more traditional deterministic approaches.
- The developed user-friendly interfaces and utilized numerical approaches make the "ANNEKs" software package useful for food scientists and engineers.

ACKNOWLEDGMENT

The author wishes to thank Dr. Sergio Almonacid, Department of Chemical and Environmental Engineering, Santa Maria Technical University, Chile, for collaboration related to modelling of food protein hydrolysis kinetic, which gave the additional opportunity to create the "ANNEKs" software.

REFERENCES

[1] Urwaye A. P. 2008. New food engineering research trends, Nova Science Publishers, Inc. New York, USA, 2008, 298.

[2] Nielsen S. S. 2010. "Food Analysis" , Springer, New York, USA, 2010, 550 p.

[3] Wei, Q., Zhimin, H. 2006. Enzymatic hydrolysis of protein: mechanism and kinetic model. Front. Chem. China, 3: 308-314.

[4] Almonacid S., Castillo C., Pinto M., Abakarov, A., Simpson R. 2008. Enzymatic hydrolysis of fish protein by-products: Kinetic analysis. Abstracts / Journal of Biotechnology 136S: S356–S401.

[5] Abakarov A., Almonacid S., Urtubia A, Simpson, R. 2010. Artificial neural network benefit in modelling of protein hydrolysis kinetic. International Conference of Food Innovation, Spain, Valencia. Book of abstract edited by Pedro Fito and fidel Toldrá, p.111.

[6] Abakarov A, Sushkov Yu, Almonacid S., Simpson R. 2009. Multi-objective optimization based on adaptive random search method: optimization of food processing. Journal of Food Science, 74(9), E471 - E487.

[7] Abakarov A., Teixeira A., Simpson R., Pinto M., Almonacid S. 2010. Modelling of squid protein hydrolysis: Artificial Neural Network approach. Journal of Food Process Engineering. DOI: 10.1111/j.1745-4530.2009.00567.

.