Chapter 9

Modelling the influence of pH and temperature on the response of an acetylcholinesterase biosensor using machine learning methods

Edwin R. García Curiel, Larysa Burtseva, Margarita Stilianova Stoytcheva, Félix F. González-Navarro, Ana Sofia Estrella Sato

Universidad Autónoma de Baja California, México.

edwin.garcia@uabc.edu.mx, burtseva@uabc.edu.mx, margarita.stoytcheva@uabc.edu.mx, fernando.gonzalez@uabc.edu.mx, sofia.estrella@uabc.edu.mx

Doi: http://dx.doi.org/10.3926/oms.189

Referencing this chapter

García Curiel, E.R, Burtseva, L., Stilianova-Stoytcheva, M., González-Navarro, F.F., & Estrella Sato, A.S. (2014). Modelling the influence of pH and Temperature on the response of an acetylcholinesterase biosensor using Machine Learning Methods. In M. Stoytcheva & J.F. Osma (Eds.). *Biosensors: Recent Advances and Mathematical Challenges*. Barcelona: España, OmniaScience. pp. 185-202.

1. Introduction

Among the chemicals that damage the environment, the ones based on the inhibitions of the enzyme acetylcholinesterase (AChE) are the predominant agriculture insecticides, carbamates or organophosphates, but they bring serious health and environment risks. The use of electrochemical biosensors is the most commonly way of detection for AChE inhibitors based on carbamates. A biosensor is a device capable to produce an electrical signal resulting from a chemical reaction between a biological compound and any other substance, producing valuable information that can be analyzed.

The induced chemical reaction (between the enzyme AChE and the neurotransmitter acetylcholine (ACh) is highly dependent of external factors, such as pH and temperature; they affect the reaction and similarly, affect the biosensor readings –i.e its accuracy and the observed electrical power. The pH takes different values, one of which shows the largest electric power and is considered as the optimum value. In the case, of temperature, the electric current has an exponential behavior, however it reaches a maximum point around 60°C, where the enzyme undergoes a denaturation process causing the current decrease until disappears.

Since the current produced by the chemical reaction depends on the interaction of many variables, it cannot be explained by simple observation; therefore a computational approach is needed in order to model the faradaic current behavior. So far, several efforts have been made to model the resulting current using different biosensors:

- Digital simulation of the current response in steady state obtained by an amperometric biosensor for a glucose system; it maintains a solid mathematical basis, but the development of the model is committed to low electrical simulations, otherwise the margin of error increases (Mell & Malloy, 1975).
- Modeling of the resulting current from an electrochemical oxygen biosensor, however it only shows the models used and do not describe how it was implemented or tested (Rangelova, Tsankova & Dimcheva, 2010).
- A chip implementation using a neural network to simulate the current on an enzymatic-biosensor; it shows the process used, but not the number of samples used to test the model, or how the experiments were performed (Alonso, Istamboulie, Ramirez-Garcia, Noguer, Marty & Muñoz, 2010).

In this work, we step aside from these traditional approaches by tanking advantage of powerful computational regression models, whose theoretical background comes from the statistical pattern recognition field.

The proposal explained trough this contribution embodies several stages briefed as follows:

- Data analysis
- Pre-processing
- Algorithm settings
- Training
- Validation

As well as its evaluation:

- Tests
- Comparative analysis

Given an electric charge produced by the interaction between the enzyme AChE and the substrate Ach, with specific temperature and pH; the regression model is selected by a comparative analysis of different regression algorithms trough a certain performance measure explained later. Under this general approach, the response of an acetylcholinesterase biosensor will be studied, and a final mathematical model will be explained.

2. The Biosensors

2.1. Electrochemical Biosensor

According to the IUPAC (International Union of Pure and Applied Chemistry) definition, a biosensor is an analytical device that combines a biological component for molecular recognition and a signal processing device called a transducer, which can detect and measure, quickly and accurately, the signal produced by the interaction of the biological element and the substance of interest. The transducer, which normally ensures high efficiency of the sensor, can be thermal, optical, magnetic, nano-mechanic, piezoelectric or electrochemical. Furthermore, the selectivity of detection is ensured by a biological recognition element, which is based on a bio-ligand deoxyribonucleic acid (DNA, RNA ribonucleic acid, antibodies, etc.), or a biocatalyst (some redox proteins, individual enzymes and enzymatic systems, such as cell membranes, complete microorganisms) (Thevenot, Tóth, Durst & Wilson, 1999). Table 1 shows several classification criteria for biosensors depending on:

- Interaction type between the components to be detected;
- Interaction detecting method;
- Biological element to recognize;
- Device transducer type.

Interaction Type	**Interaction Detecting Method**
Biocatalitic Bioafinity	Direct Indirect
Biological Element to Recognize	**Transducer Type**
Enzyme Organelle, tissue or complete cell Biological receptor Antibody Nucleic acid Aptamers	Electrochemical Optic Piezoelectric Thermometric Nanomechanic

Table 1. Biosensors classification criteria

The electrochemical biosensors are divided into two types:

- Potentiometric.
- Amperometric.

Biosensors that measure only a change in potential at the interface-analyte sensor with respect to the reference electrode are known as potentiometric sensors or biosensors. Sensors that impose external potential to effect the transformation are referred to as electrochemical *amperometric biosensors.* Inhibition of the enzyme in these biosensors is monitored by the current change, detecting to a certain potential oxidation or reduction. Table 2 describes various environmental biosensors according to the type of transducer used on certain biological elements, compound to detect and the area possibly analyzed.

Transducer type	**Biological element**	**Analite**	**Area**
Electrochemical	Antibodies	Atrazine	*
Optic	Antibodies	Simazine	Ground, water mass
Optic	Antibodies	Pesticides	Rivers
Electrochemical	Antibodies	Tensoactives (alkylfenols)	*
Electrochemical	Antibodies	Estradiol	*
Electrochemical	Antibodies	Eschenchacoll	Drinking water
Optic	Antibodies	Enteritis Listeria monocytogenes Salmonella	*
Acoustic	Antibodies	Salmonella typhymurium	*
Optic	Enzyme (AChE)	Organophosphates compounds	Water
Electrochemical	Enzyme (AChE)	Paraoxon y carbofuran (pesticides)	Residual waters
Electrochemical	Enzyme (tyrosinase)	Fenols	Ground, dirt, waters

Table 2. Biosensors with environmental applications

2.2. Enzyme Reaction

The measurements were performed using an electrochemical biosensor with two electrodes, a counter electrode and a reference electrode in conjunction with an insulated electrode (working electrode), which in its lower part, has a portion (mg/cm^3) of the enzyme AChE. Figure 1 shows the schematic AChE biosensor used in the experiment.

The biosensor (working electrode) used in the experiment was prepared by immobilizing the AChE enzyme, chemically bonded on the surface of a graphite electrode. The auxiliary electrode was glassy-carbon type and the reference electrode a saturated calomel.

Thereafter the electrodes are immersed into a conventional electrochemical cell, which contains a buffer solution with a given concentration (mmols/L) of the substrate ACh. The solution receives a fixed acidity (pH) amount and temperature (˚C). During the study the electrode maintained a constant rotation speed of 1000 rpm.

Equations of chemical reactions where the AChE is involved are shown below:

$$\xrightarrow{\text{AChE}}$$

$$ACh + H_20 \Rightarrow Ch + \text{Acetic Acid} \quad (1)$$

$$2Ch = Ch\text{-}Ch + 2e^- + 2H^+ \quad (2)$$

During the study we measured the current produced by the oxidation reaction product Ch (Equation 2) to the enzymatic reaction (Equation 1).

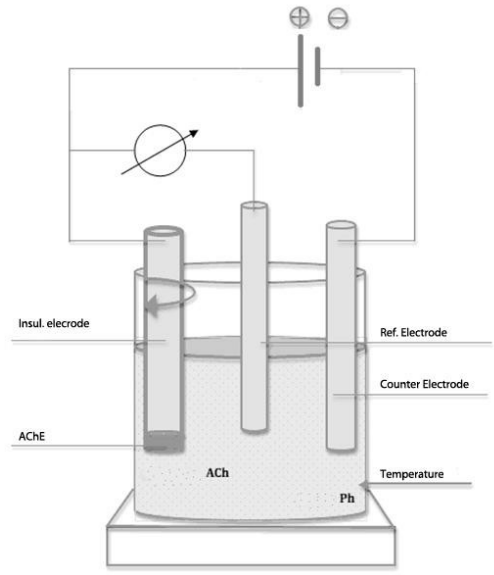


Figure 1. Acetilcholinesterase biosensor scheme

2.3. Influencing Factors

2.3.1. Substrate concentration

The behavior of the current change with the concentration of ACh is total. At higher concentrations, better results are obtained depending on the current. Current function I vs. Cs is described by a hyperbolic function:

$$I = Imax * Cs / (Cs + Km), \quad (3)$$

where Km is a called Michaelis-Menten constant. Thus, the optimum concentration of ACh is the highest level used: 1.0 mM / L. Table 3 shows the typical experimental results.

pH	Steady-state current, µA					
	25˚C	**30˚C**	**40˚C**	**50˚C**	**60˚C**	**70˚C**
5	8.80	14.06	15.67	17.27	18.88	0.72
6	20.30	32.44	36.14	39.14	43.56	1.66
7	21.90	35.00	39.00	43.00	47.00	1.80
8	17.90	25.09	31.87	35.14	38.41	1.47
9	15.70	25.09	27.95	30.82	33.69	1.29

Table 3. Optimal results as a function of ACh concentration fixed at 1.0 mM/L (pH 5-9 and temperature 25-70˚C)

2.3.2. pH

The response of the steady-state current, to the increase of the acidity levels, behaved in ascending form. The optimum pH greatly varies depending on temperature and Cs changes. If Cs ≤ 0.6 mM/L, and t ≤ 30˚C, the optimum pH is found at level 7; if the t > 30˚C, then it is found at level 8. If Cs > 0.6 mM / L, the optimum pH is found at level 7.The obtained data is presented in Table 4.

Substrate Concentration, mM/L	Steady-state current, µA				
	pH 5	pH 6	pH 7	pH 8	pH 9
0.2	8.56	11.22	10.65	15.79	13.13
0.4	13.32	22.27	21.32	28.36	24.55
0.6	15.68	30.98	31.98	32.35	28.62
0.8	18.11	39.85	42.63	36.87	32.39
1.0	18.88	43.56	47.00	38.41	33.69

Table 4. Optimal results in terms of pH and Cs, with a preset temperature at 60˚C

2.3.3. Temperature

The response of the steady-state current to the increase in temperature is upward from 4.5 µA to 47 µA. The latter value is the maximum current found at a temperature of 60 ˚C. When the temperature goes from 60 to 70˚C, the current decreases drastically, such that the value obtained at maximum temperature (0.382075µA, found at 70˚C, pH 7, 0.2 Cs) is lower than the found at 25˚C (minimum temperature) (Table 5).

Substrate Concentration, mM/L	Steady-state current, µA					
	25˚C	30˚C	40˚C	50˚C	60˚C	70˚C
0.2	5.60	7.63	8.68	9.60	10.65	0.38
0.4	11.20	15.26	17.37	19.20	21.31	0.76
0.6	16.30	22.89	26.06	28.80	31.97	1.14
0.8	20.00	30.52	34.75	38.40	42.63	1.52
1.0	21.90	35.00	39.00	43.00	47.00	1.80

Table 5. Optimum results in terms of temperature and Cs, with fixed pH at level 7

3. Machine Learning Algorithms

3.1. Neural Networks

Neural networks are computational models based on the structure of nerve cell connections found in the brain of living things, and likewise try to imitate their operation (Russell & Norvig, 2010; Wolfgang, 2011; Luger, 2009; Mitchell, 1997).

A neural network is composed by one or more units (neurons), which possess a series of connections serving as inputs (dendrites), outputs (axioms), or to communicate with another unit (synapses). Each connection has an associated numerical weight given. A neuron processes the input information and produces an output, which could be considered as an input to another neuron or as a final result.

A neural network operates in the following way. A vector $x = (x_1, \ldots, x_n)$, where $x_j \in R$, $j = 1,\ldots,n$, is used as input. A numerical weight w_{ji}, where i is the index of neuron, is associated with each element x_j. These values are then summed on each neuron y, and it is applied a transfer function (activation) f, which determines the output value y_i:

$$y_i = f\left(\sum_{j=1}^{n}\left(w_{ji}x_j\right)\right) \quad (4)$$

The activation function most often used are the hyperbolic tangent function (5) or a sigmoid function (6):

$$\tanh(x) = \frac{1-e^{x}}{1+e^{x}} \tag{5}$$

$$sig(x) = \frac{1}{1+e^{x}} \tag{6}$$

There are different neural network architectures, developed for different types of problems. If the problem is to model (regression) or predict (classification) data, then the architectures are used as a single layer perceptron (SLP), for linearly separable problems, or multilayer perceptron (MLP), for high-dimensioned problems or not linearly separable. In addition, these architectures are based on different learning algorithms, such as the Least-Mean-Square (LMS) algorithm or the back-propagation error algorithm (BP), considered as the learning algorithm used by MLP type networks.

The MLP-BP is one of the most solicited neural network architectures in regression tasks, given its adaptability to different problems. In this case, the regression problem of approximating a possible nonlinear function $f(x)$ with a neural network $Y(x)$ BP MLP architecture, where $x \in R^n$. Figure 2 shows the scheme of a network MLP-BP.

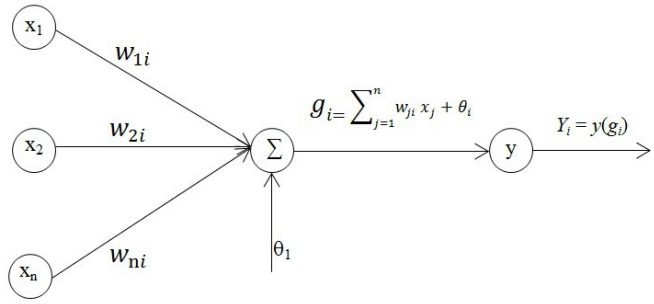


Figure 2. Example of a net MLP-BP

Inputs x_j, j = 1... n, of the neuron i are multiplied by the weights w_{ji} and are summed with the bias value θ_i.. The result g_i is the input to the activation function Y_i. The output node i becomes:

$$Y_i = y\left(\sum_{j=1}^{n} w_{ji} x_i + \theta_i\right) \tag{7}$$

3.2. Support Vector Machines

The support vector machines (SVM) are supervised learning methods that generate a mapping function from a set of pre-labeled training data. The mapping can be either a classification or a regression function.

In a classification problem, a mapping is used to transform the input data, which is not linearly separable in the original space, to a space of greater dimension, which happens to be separable. The produced model depends only on a subset of the original input data whose characteristic is that it creates a boundary that separate one data class from the others. The data belonging to

that subset are called *support vectors*. This boundary named margin assure us a maximum distance between the different classes (Figure 3).

The goal of the SVM is to create a computational model to predict the class label of new data samples.

In addition to their solid mathematical foundation based on statistical learning theory, the SVMs have shown a highly competitive performance in many applications such as bioinformatics, text mining, face recognition, image processing, which have been established the SVM as one of the basic tools of machine learning.

A regression problem is expressed in terms of the MSV as follows:

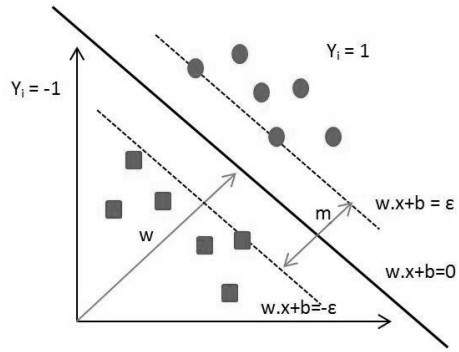


Figure 3. Classification of linearly separable data

There are given a training data set **D** = {x_i, y_i}, i = 1… l, of input vectors $x_i \in \mathbf{R}^n$ and their labels $y_i \in \mathbf{R}^1$. Through parameter C > 0 and $\varepsilon > 0$, the standard form of the SVM applied to a regression problem is (Vapnik, 1998; Chang & Lin, 2011):

$$min_{\alpha,\alpha^*} \frac{1}{2}(\alpha-\alpha^*)^T Q(\alpha-\alpha^*)+\varepsilon\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{l} y_i(\alpha_i-\alpha_i^*) \tag{8}$$

s. t.

$$e^T(\alpha-\alpha^*)=0, \tag{9}$$

$$0\le\alpha_i,\alpha_i^*\le C, i=1,\dots,l, \tag{10}$$

where

$$Q_{ij}=K(x_i,x_m)\equiv\phi(x_i)^T\phi(x_m) \tag{11}$$

After solving the problem (8), the approximation function is:

$$g_{(x)}=\sum_{i=1}^{l}(-\alpha_i+\alpha_i^*)K(x_i,x)+b \tag{12}$$

where ε is a predefined constant that controls the noise tolerance. With the insensitive loss function ε, the objective is to find the function $g_{(x)}$ whose deviation is (at most) the value of the

loss function ε from obtained etiquettes y_i for all training data, which should be as flat as possible.

In other words, the regression algorithm does not affect errors as long as they are less than ε, but any deviation greater than ε is not accepted.

The defined function:

$$K(x_i, x_m) \equiv \phi(x_i)^T (x_m) \tag{13}$$

is called kernel function (Figure 4). Although researchers frequently propose new kernel functions, we suggest the use of the following (Pedroza, 2007; Hammel, 2009; Hsu, Chang, & Lin, 2003; Chang & Lin, 2011):

- Linear: $K(x_i, x_m) = x_i^T x_m$.
- Polynomial: $K(x_i, x_m) = (x_i^T x_m + r)^d$, d > 0.
- Radial Basis Function: $K(x_i, x_m) = \exp(-\gamma \| x_i - x_m \|^2), \gamma > 0$.
- Sigmoidal: $K(x_i, x_m) = \tanh(\gamma x_i^T x_m + r)$.

Note: *r* and *d* are specific parameters of the functions.

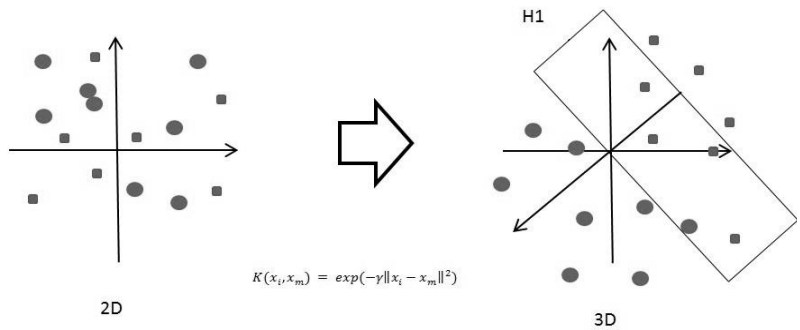


Figure 4. Transformation scheme from a set of data using a radial basis function kernel

The kernel function that reaches the objectives previously described is not known a priori; hence the procedure for its finding has not been established. The way to meet the function ϕ is based on the performance of different tests on known ϕ functions. This procedure is called *grid search*; some functions (polynomial, radial, sigmoidal, etc.), require specific parameters, which determine the separability between classes. The consumption of computational resources depends on the dimensions of the grid search (amount of features to be compared and the amount of different parameters for each function to be used), as well as the size of the data set.

Although there are different kernel functions, it is common the use of Radial Basis Function. However it can be chosen using other kernels depending on the results obtained for a particular case.

3.3. Data PreProcessing

The 150 samples measured on the electrochemical experiment were analyzed, concluding that the differences between the ranges of the parameters (Cs, pH, T) might influence the training

algorithms. For this reason a min-max [0, 1] normalization process was applied. Table 6 shows the values of the data before and after standarization.

Parameters	Original Data	Normalized Data
Cs	[0.2, 0.4, 0.6, 0.8, 1.0]	[0.0, 0.25, 0.50, 0.75, 1.0]
pH	[5, 6, 7, 8, 9]	[0.0, 0.25, 0.50, 0.75, 1.0]
T	[25, 30, 40, 50, 60, 70]	[0.0, 0.1111, 0.3333, 0.5556, 0.778 ,1.0]

Table 6. Original and Normalized Data Ranges

After normalizing the data, we proceed to divide the entire set into subsets for training, validation and testing. Because the entire data set is considered small (150 samples), it is necessary to use a resampling technique, which will also permit to separate the main assembly into three subsets of data instead of 2. The first is a subset assigned for training-validation and the second is used only for testing. The resampling technique chosen is a version of the K-Fold CV, which is repeated k-times, in this way, a larger number of samples can be trained, and obtain a better generalization error. Furthermore, this technique allows the model to be evaluated in a better way because more elements are considered for validation, and reinforces the model's behavior against "new" samples.

3.4. The Regression Model

3.4.1. ANN: Training and Validation

In order to find the neural network parameters that render the Means Square Error --i.e. no. of neurons per layer, no. of internal layers, learning rate and learning algorithm, a grid search process was carried out. Considering that the search process was long, the combination of parameters for the algorithm calibration has been reduced based on previous works about approximations to functions of biological characteristics or obtained through the use of biosensors (Hsu et al, 2003).

ANN Configuration:

Layers	5
Neurons	30 per each layer
Training	Levenberg-Marquadt Algorithm
Learning	Gradient Descendent
Resampling technique	*k* x *k*-Fold CV, *k* ϵ (5, 10)
Performance measure	Mean Square Error (MSE)

The presented configuration offered the best results in terms of the time consumption, resources and good performance measure during training. Other prominent configurations improve the network performance on an insignificant manner; however because of their trend to consume considerable amounts resources, they were discarded. The selected configuration is programmed with 30 neurons in each of the five layers. Levenberg-Marquadt Algorithm was used to optimize the learning process. To improve the evaluation of the model a K x K-Fold CV resampling technique (Refaeilzadeh, Tang & Liu, 2009) with two different values of K (5, 10) was used.

3.4.2. SVM: Training and Validation

After the preprocessing and data division, a grid search process was applied using an exponential increasing of the parameters, as it is suggested in (Hsu et al., 2003), aiming to find the best SVR algorithm configuration in the shorter time. Since the performance of this algorithm depends on the kernel function employed, several searches were conducted, one for each type of kernel: linear, polynomial, radial basis and sigmoid. The best results were showed with a radial basis kernel function. Table 7 shows the best results of each of the 4 kernels.

The search process using the radial basis kernel was performed on 3 parameters: the error penalty (C), insensitive loss function (ε) and the radial kernel parameter (α). The resampling technique was used in this step.

	Linear	**Polynomial**	**Radial Base Function**	**Sigmoid**
Configuration	c = 29 ε = 22.21	c = 29 ε = 22.21 α = 0.075 d = 2	c = 512 ε = 4.6268 α = 0.075	c = 29 ε = 22.21
MSE, 5-Fold CV	40.53	26.75	5.41	7.83
MSE, 10-Fold CV	84.72	53.78	10.84	15.67

Table 7. Best SVR kernel results

The performance comparison between the algorithms mentioned above ensures that the configuration presented is the best possible for the SVR algorithm.

SVM configuration:

C	$2^9 = 512$
α	$2^{2.21} = 4.6268$
ε	0.075
Kernel	Radial Base Function
Resampling technique	k x k-Fold CV, $k \in (5, 10)$
Performance measure	MSE

3.3. Neural Networks

The tests on the ANN-MLP model are realized to make a simulation of CS parameters, pH and T using the test data set and the ANN-MLP model selected during Grid Search. Figure 5 shows a comparison of the original data and the data predicted by the model ANN-MLP. A simulation of the original data and the data predicted by the ANN-MLP is exposed on Figures 6a and 6b, respectively.

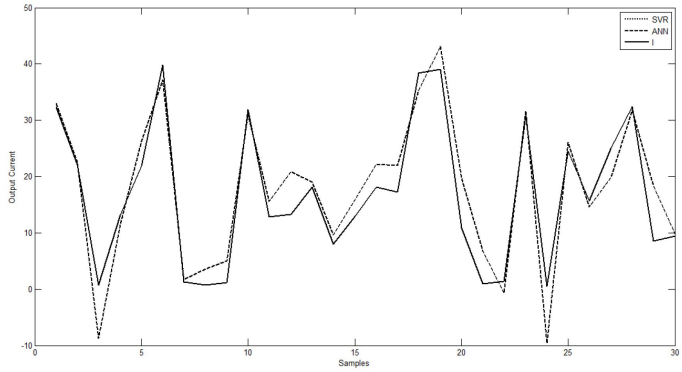
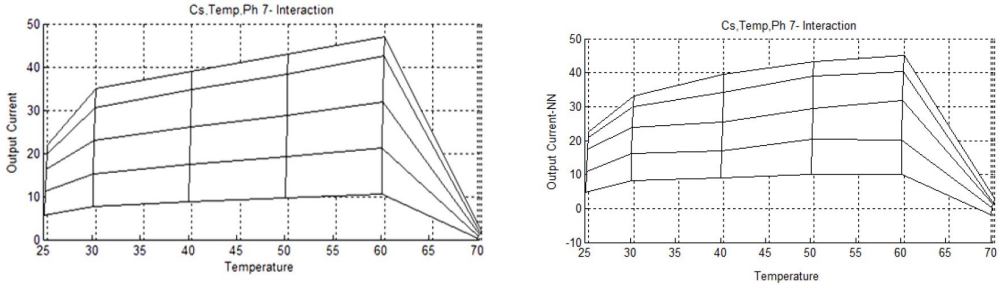


Figure 5. Graphs of the resulting current test data vs. the data predicted by the ANN-MLP model



(a) (b)

Figure 6. Comparative graphs of the ANN-MLP simulated data and the experimental data: (a) Experimental data, (b) ANN-MLP predicted data

3.4. SVM

Similarly to the ANN-MLP, the test on the SVR model consist of the prediction of the test data set, using the best configuration found by means of a grid search process; and then, compare this predicted data set with the original test data. Table 8 shows a part of the comparison of test data and simulation using the SVR model.

No	Test Data *I*, μA	Simulation *I*, μA	Error
1	32.35	31.7865191	0.56348089
2	22.11	22.8484843	0.73848435
3	0.72	-3.47229928	4.19229928
4	13.13	14.0527819	0.9227819
5	21.9	24.0532271	2.1532271
6	39.85	40.5384729	0.68847285
7	1.32	1.76999286	0.44999286
8	0.74	3.11594287	2.37594287
9	1.16	1.50589061	0.34589061

Table 8. SVR Model best results

The comparison of SVR against the test output parameters from the original data is shown on Figure 7. A simulation of the original data and the data predicted by the SVR is exposed on Figures 8a and 8b, respectively.

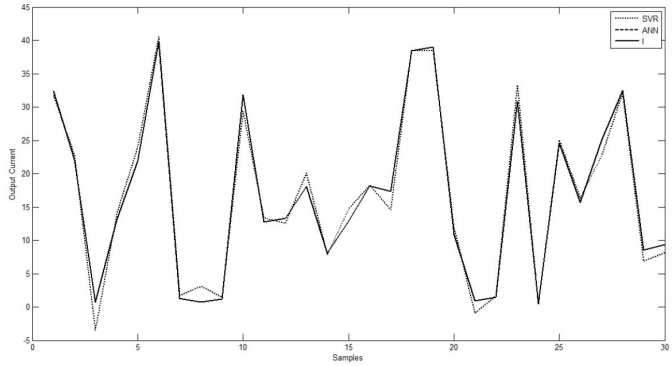
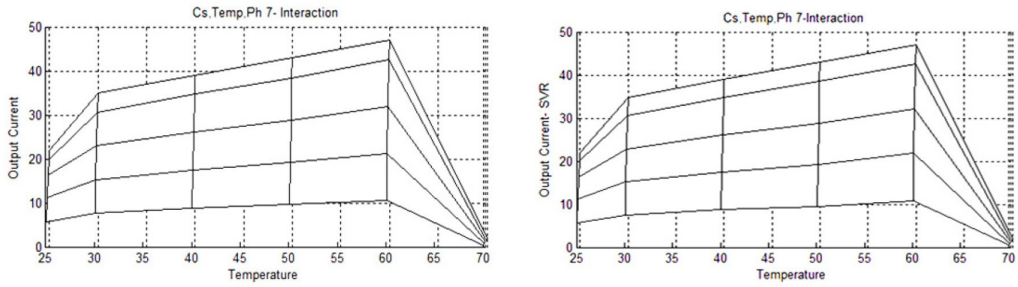


Figure 7. Current data resulting from the test set vs. the data predicted by the SVR



(a) (b)

Figure 8. Comparative plots of the SVR simulated data and the experimental data: (a) Experimental data, (b) SVR predicted data.

A comparative analysis was realized by observing the graphs of the test data modeled by the algorithms, as well as comparing errors that have the simulations of the test data and the MSE that each model has in total. Figure 9 shows the experimental test data and the data simulated by ANN-MLP and SVR models.

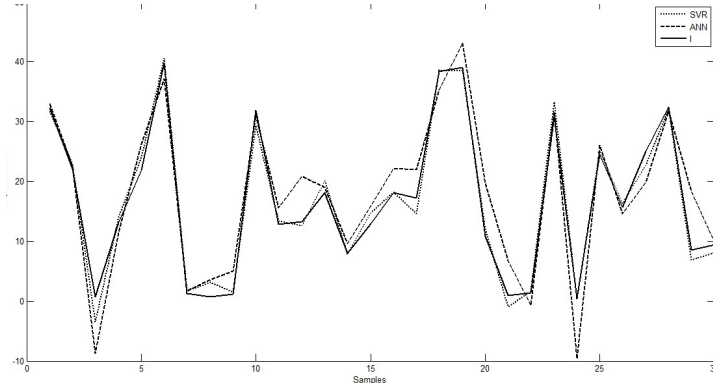


Figure 9. Comparison of experimental test data vs. data predicted by the ANN-MLP and SVR models

Table 9 shows the quantitative comparison between the test data simulated by the models and experimental data, based on the MSE of each model, 3 samples were taken randomly from the set of test data. Figure 9 and Table 9 show that both models provide good approximation function, but the SVR is the best fitting model, so it is considered the best.

No	Test Data Samples				ANN MLP		SVR	
	C_s, mmol	pH	*t*, ˚C	*I*, µA	*I*, µA	Error,	*I*, µA	Error,
1	0.2	6	30	8.03	9.68	1.65	7.84	0.19
2	0.8	7	50	18.15	22.21	4.06	18.26	0.09
3	0.4	7	40	38.41	35.27	3.14	38.56	0.15
Error (MSE), %					-	2.95	-	0.143

Table 9. Models quantitative comparison

In order to develop a simulation using more realistic features, a new set of samples has been generated, applying the minimum measurable level of ACh, pH and t. (Table 10).

Feature	Scale	Increment
Substrate Concentration	$0.2 \leq C_s \geq 1.0$	0.01
pH	$5 \leq pH \geq 9$	0.01
Temperature	$25 \leq t \geq 70$	0.1

Table 10. Scale simulation features

The samples generation process should consider the following:

- The data range of new samples should be the same as that of the original data.
- The generated samples should be normalized in the same way that the experimental data, using the same set of equations.

This must be hold as a way to maintain a defined structure, and avoid data extrapolation problems, causing a poor spread of new samples and therefore a higher error rate in the simulation.

Recalling that the number of samples used to develop the model were 150, with the new simulation levels, 15 467 031 samples were generated. However, in previous sections it was shown the dependency of the response variable with the Cs value; at higher concentrations, the amperage was increased and vice versa. In this way, and in order that the simulation can be analyzed qualitatively, the process was made from the perspective of T and pH, leaving the Cs value at 1.0 mM/L. For this reason, the samples were reduced to 225, 951, decreasing the samples generation, normalization and prediction time of the regression model.

To evaluate the simulation, a comparative analysis between simulated samples and experimental data were performed. Figures 10 and 11 show the minimum scale feature simulation, with Cs = 1.0 mM/ L. The results show that the function keeps the behavior of the experimental data with minor irregularities which do not compromise their performance.

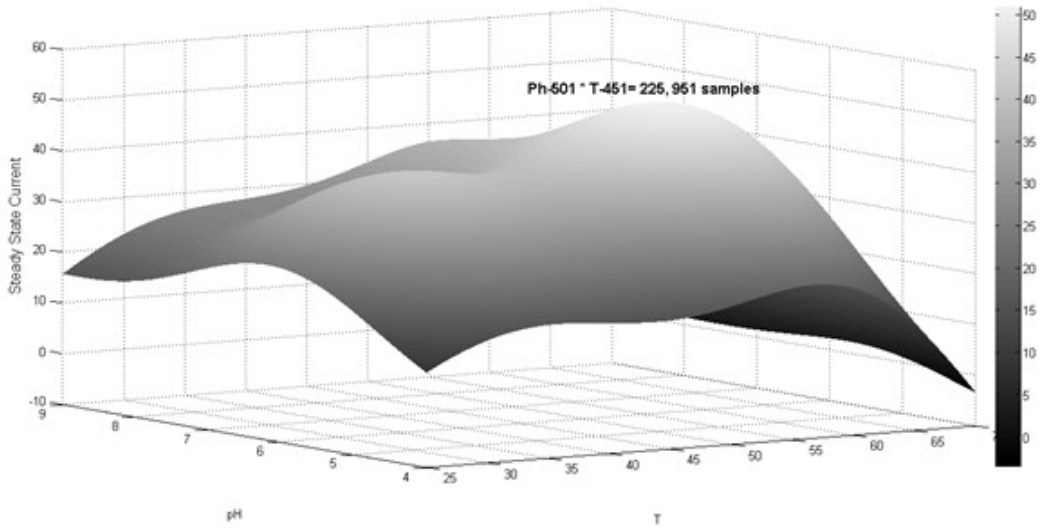


Figure 10. Minimum scale feature simulation

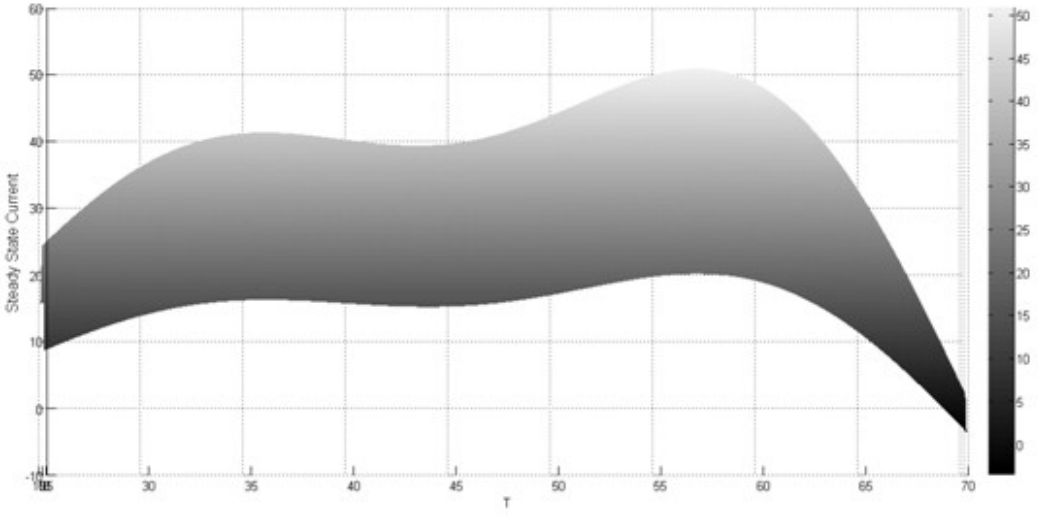


Figure 11. Comparison between original and simulated data: Temperature T perspective

In order to find the sample that generates the highest response value, a search through the whole simulation samples was made. Using this simulation, intervals for locating the maximum response value were generated; shown on Table 11.

Interval (μA)	Temperature [0.1]	pH [0.01]	# of samples in the intervals
±0.001	56.9	6.03, 6.04	2
±0.01	56.8≤ t ≤ 57.1	6.00≤ pH ≤ 6.07	24
±0.02	56.7≤ t ≤ 57.2	5.99≤ pH ≤ 6.08	48
±0.05	56.5≤ t ≤ 57.3	5.96≤ pH ≤ 6.11	117
±0.1	56.3≤ t ≤ 57.5	5.93≤ pH ≤ 6.14	232
±0.2	56.0≤ t ≤ 57.8	5.88≤ pH ≤ 6.19	472
±0.5	55.5≤ t ≤ 58.3	5.79≤ pH ≤ 6.29	1,117
±1.0	54.8≤ t ≤ 58.8	5.69≤ pH ≤ 6.41	2353
±2.0	53.8≤ t ≤ 59.6	5.55≤ pH ≤ 6.58	4810
±5.0	51.40≤ t ≤ 61.1	5.28≤ pH ≤ 6.98	12,896

Table 11. Optimum interval location

The two samples presented in the range ± 0.001 locate the maximum value of the response variable to a detectable level, and the maximum is achieved with the combination:

C_s = 1.0 mM/ L, **T** = 56.9°C, pH = 6.03,

generating a steady-state current of 50.9502 μA, and the second sample that have the same configuration with the exception of pH(6.04), generates approximately the same current, with a difference of 0.000004 μA. Figure 12 highlights the area within the range ± 1.0 μA, and locates the maximum current value expressed through the simulation of samples (Garcia, Burtseva, Stoytcheva & Gonzales, 2011).

4. Conclusions

This research resolves the doubts found in the literature about learning of biological functions provided by the use of electrochemical sensors. Keeping the virtues of these, a detailed process is presented in the development of different learning models, and the procedure for the evaluation of results.

Experimental data are analyzed through different perspectives (Cs, pH, T), also their behavior and location of the possible areas of the extreme values in the resulting current, with the aim of finding the combination of parameters that maximize the sensitivity of the determination of ACh.

According to the analysis of the state of the art, the Artificial Neural Networks are the most suitable regression models to the task at hand, i.e. to predict and analyze the parameters of a biosensor. In the same way, the SVMs offer a solid and competitive performance with the possible advantage that it has less tuning parameters than a Neural Network. Although the results obtained using Neural Networks were satisfactory, the SVMs show the best performance, i.e. function approximation, both in testing and simulation modeling process.

It must be noticed that educated practices in machine learning literature instruct to computer scientists to have a proper independent test data set as a way to correctly assess the generalization capacity of any model; therefore, new test data samples must be available.

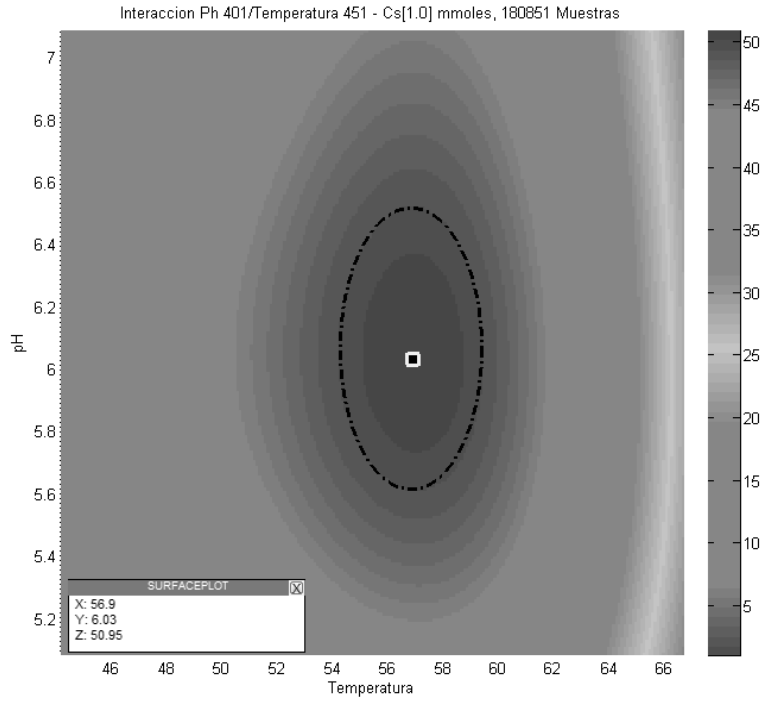


Figure 12. Location of the maximum current value

References

Alonso, G., Istamboulie, G., Ramirez-Garcia, A., Noguer, T., Marty, J., & Muñoz, R., (2010). Artificial neural network implementation in single low-cost chip for the detection of insecticides by modeling of screen-printed enzymatic sensor response. *Computers and Electronics in Agriculture, 74*, 223-229. http://dx.doi.org/10.1016/j.compag.2010.08.003

Chang, Ch.-Ch. & Lin Ch.-J., (2011). *LIBSVM: A Library for Support Vector Machines.* http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf (Last access date: March 2012).

Garcia, E.R., Burtseva, L., Stoytcheva, M., & Gonzales, F.F., (2011). Predicting the behavior of the interaction of acetylthiocholine, pH and temperature of an acetylcholinesterase sensor. *LNCS-LNAI, 7094*, 583–591.

Hammel, L. (2009). *Knowledge discovery with support vector machines.* A John Wiley & Sons, Inc., Hoboken, NJ, USA, 246P.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A practical guide to support vector classification.* Technical report, Department of Computer Science, National Taiwan University. http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf (Last access date: March 2012).

Luger, G.F., (2009). *Artificial Intelligence: structures and strategies for complex problem solving. Pearson Addison-Wesley,* 754.

Mell, L.D., & Malloy, J.T., (1975). A Model for the Amperometric Enzyme Electrode Obtained through Digital Simulation and Applied to the Immobilized Glucose Oxidase System. *Analytical Chemistry, 47*, 299-307. http://dx.doi.org/10.1021/ac60352a006

Mitchel, T.M., (1997). *Machine Learning.* McGraw-Hill Science, 432.

Pedroza, G., (2007). Aplicación de las máquinas de soporte vectorial a reconocimiento de hablantes. Universidad Autónoma Metropolitana. México. http://cbi.izt.uam.mx/foroacademico/2007/res/cartel27.pdf (Last access date: January 2013).

Rangelova, V., Tsankova, D., & Dimcheva, N., (2010). Soft Computing Techniques in Modeling the pH and Temperature on Dopamine Biosensor. In Vermon S. Somerset (Ed.). *Intelligent and Biosensors.* Croatia: InTech. 99-122. http://dx.doi.org/10.5772/7029

Refaeilzadeh, P., Tang, L., & Liu, H., (2009). *Cross-Validation. Encyclopedia of Database Systems.* Springer, 532-538.

Russell, S.J., & Norvig, P., (2010). *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1132.

Thévenot, D.R., Tóth, K., Durst, R.A., &Wilson, G.S., (1999). Electrochemical biosensors: recommended definitions and classification. *Pure and Applied Chemistry, 71(12)*, 2333-2348. http://dx.doi.org/10.1351/pac199971122333

Vapnik, V., (1998). *Statistical Learning Theory.* New York: Wiley. 736.

Wolfgang, E., (2011). *Introduction to Artificial Intelligence.* Springer. 316.