

# Predictive Analysis of Blood Gasometry Parameters Related to the Infants Respiration Insufficiency

Wieslaw Wajs<sup>1</sup>, Mariusz Swiecicki<sup>2</sup>, Piotr Wais<sup>1</sup>, Hubert Wojtowicz<sup>1</sup>,  
Pawel Janik<sup>3</sup>, and Leszek Nowak<sup>4</sup>

<sup>1</sup> University of Mining and Metallurgy, Institute of Automatics,  
Cracow, Poland

<sup>2</sup> University of Technology, Intitute of Computer Modelling,  
Cracow, Poland

<sup>3</sup> Jagiellonian University, Faculty of Mathematics,  
Astronomy and Applied Informatics,  
Cracow, Poland

<sup>4</sup> Jagiellonian University, Faculty of Mathematics and Informatics,  
Cracow, Poland

**Abstract.** The paper presents application of artificial immune system in time series prediction of the medical data. Prediction mechanism used in the work is basing on the paradigm stating that in the immune system during the response there exist not only antigene - antibody connections but also antigene - antigene connections, which role is control of antibodies activity. Moreover in the work learning mechanism of the artificial immune network, and results of carried out tests are presented.

## 1 Introduction

Biological phenomena, and particularly organic processes are of a dynamic character. Diagnosis and therapy are of the same character. In approaching diagnosis and making therapeutic decision phenomena happening in time are investigated [1]. Every doctor knows that it is rare to make a decision about diagnosis and treatment on the basis of only one clinical observation. Usually it is based on several patient's examinations, regular analysis of many biophysical and biochemical parameters, or imaging examination. On a particular level of these examinations preliminary diagnosis is determined, and later a final one. Basing on determined diagnosis, and on the evaluation of the sickness process dynamics, decision about the treatment is taken. The field of neonatology has made tremendous progress during the last twenty years. It's been made possible thanks to the progress in modern technologies of intensive medical care and progress in the applied sciences. Diagnostic examinations are based on micro - methods. Modern monitoring technologies, both non - invasive and invasive, aren't very strenuous for the patient, allow in an exact manner for continuous monitoring of gas exchange, lungs mechanics and functioning of circulatory system. Patophysiology of the infant respiration insufficiency and the role of the

surfactant deficiency in various forms of the respiration insufficiency was discovered. Physiology and patophysiology of the water - electrolyte economy of regular and premature neonates was also discovered. Initial stabilization of the infants state is a difficult task. It can take even few days. To achieve it the doctor analyses repeatedly many parameters related to the patient's health condition. To these parameters belong birth anamnesis, physiological examinations (body's weight, diuresis), results of additional examinations ( biochemical, micro biological, imaging) and readings from the monitoring instruments ( pulseximeter, cardio monitor, invasive measurement of the arterial pressure, respiratory mechanics monitor). It isn't rare when doctor appraises simultaneously over fifty variables. Analysis of this huge amount of data is hard and requires experience, all the more if the decision about the treatment should be made quickly. In the result of carried out data analysis doctor makes decision about the treatment expecting positive results, which are expressed by the desired changes in the results of additional examinations, readings of the monitoring instruments and physical examination. The whole process can be verified by comparing it to the model of respiratory insufficiency progress carried out by the doctor. Creation of this model is based on theoretical and empirical knowledge found in scientific books and papers and also on the doctor's experience. Out of necessity created model is of subjective character. Moreover results of analysis carried out by the doctor are always dependant of many external factors - recently read reports, weariness of the doctor etc. Intensification of gas exchange disorders are best reflected by arterial blood gasometry parameters examined in the context of currently used methods of ventilation support. For that reason as an output values four directly measured parameters of arterial blood gasometry:  $pH$ ,  $pCO_2$ ,  $pO_2$ ,  $HCO_3$  and oxygen parameter ( $pO_2/FiO_2$ ) were chosen. Alkalis deficiency and hemoglobin oxygen saturation were omitted as derivatives of the four parameters mentioned above.

### 1.1 Parameters Prognosis

Below short characteristic of the parameters is presented:

$pH$  - blood reaction is a parameter, which depends on the state of the respiratory system, but also on the circulatory system functioning, kidneys and metabolism of system. It is a general parameter, which can be treated as an indicator of the general state of the patient.

$pCO_2$  - partial pressure of the carbon dioxide is strictly correlated with the rate of the respiratory system efficiency. Above all it is dependent on minute ventilation. Its increase causes simultaneous decrease of  $pCO_2$ . Increase of the partial pressure of the carbon dioxide is usually the symptom of respiration disorders with lowered ventilation. However in serious respiration disorders with handicapped gas diffusion through the alveolar - capillary barrier also appear symptoms of handicapped exchange of carbon dioxide. Increase of  $CO_2$  decreases blood's  $pH$ .  $pCO_2$  of patients with supported ventilation depends mainly on the frequency of breaths (RR) and respiration capacity (TV).

$pO_2$  - partial pressure of oxygen is also strictly correlated with the rate of the respiratory system efficiency. Normal exchange of oxygen in lungs depends mainly on the proper ratio of ventilation to perfusion and on permeability of the alveolar - capillary barrier. Perturbation of the ventilation to perfusion ratio is observed usually in the case of atelectasis. Thickening of the alveolar - capillary barrier in the first place handicaps oxygen exchange, only further decrease of its permeability handicaps  $CO_2$  exchange.  $pO_2$  of patients with ventilation support depends on the oxygen concentration in the respiratory mixture ( $FiO_2$ ) and on the rate of lung's dilation. The latter in the patients with ventilation support depends on the proper pressure in respiratory tracts (peak pressure of inspiration, average pressure in respiratory tracts). Oxygen deficiency in the organism causes increase of anaerobic metabolism and also increase of lactic acid and decrease of blood's pH production.

$HCO_3$  - hydrogen carbonate concentration in blood depends on system's metabolisms and kidney's functioning. In case of disproportions between transport of oxygen to tissues and metabolic requirements of system decrease of bicarbonate concentration and metabolic acidosis supervene.

Assessing particular clinical case one doesn't rely on only one, even complex parameter, usually one evaluates many various parameters. Because of that four parameters were chosen as inputs of artificial immune network. Exact classification of their values reflects direction and rate of changes - improvement or deterioration of respiratory system functioning, occurring in the natural progress of the illness and used treatment methods.

## 1.2 Source of Data

Research data used in the presented work, was gathered during few years long observation and hospitalization of patients on the Infant Intensive Care Ward of the Polish - American Institute of Pediatrics Collegium Medicum Jagiellonian University of Cracow. On this ward a computer database is used called Neonatal Information System (NIS), which stores information about all hospitalized infants.

## 2 An Artificial Immune Network for Signal Analysis

Main goal of our system is a prediction of signals and to achieve this goal several problems had to be solved. The first of them is connected with an algorithm of learning the immune network. The next problem is related to the structures of data, which are responsible for representation of signals. Solutions to these problems are presented below. The last paragraph presents results of signals prediction by our immune network.

### 2.1 Signal Representation

The input signal for the system, is interpreted as an antibody (Ab) so the task of immune network is to find an antigen Ag that will be suitable for Ab. The Ag-Ab

representation will partially determine which distance measure shall be used to calculate their degree of interaction. Mathematically, the generalized shape of a molecule (m), either an antibody (Ab) or an antigen (Ag), can be represented by a set of real-valued coordinates  $m = \langle m_1, m_2, \dots, m_L \rangle$ , which can be regarded as a point in an L-dimensional real-valued space.

$$D = \sqrt{\sum_{i=1}^L (ab_i - ag_i)^2} \tag{1}$$

The affinity between an antigen and an antibody is related to their distance that can be estimated via any distance measure between two vectors, for example the Euclidean or the Manhattan distance. If the coordinates of an antibody are given by  $\langle ab_1, ab_2, \dots, ab_L \rangle$  and the coordinates of an antigen are given by  $\langle ag_1, ag_2, \dots, ag_L \rangle$ , then the distance (D) between them is presented in equation (1), which uses real-valued coordinates. The measure distances are called Euclidean shape-spaces.

## 2.2 Learning Algorithm

In this article, we are basing on an algorithm that was proposed in the papers [2, 3] by de Castro and Von Zuben. Our modified version of immune net algorithm [4] was adapted to cope with continuous signals prediction. The aforementioned algorithm is presented below. The learning algorithm lets building of a set that recognizes and represents the data structural organization. The more specific the antibodies, the less parsimonious the network (low compression rate), whilst the more generalist the antibodies, the more parsimonious the network with relation to the number of antibodies. The suppression threshold controls the specificity level of the antibodies, the clustering accuracy and network plasticity.

For each  $Ag_j \in Ag$  do

- Determine its affinity  $f_{i,j}$ ,  $i = 1, \dots, N$ , to all  $Ab_i$ .  $f_{i,j} = 1/D_{i,j}$ ,  $i = 1, \dots, N$
- A subset  $Ab(n)$  which contains the n highest affinity antibodies is selected;
- The n selected antibodies are going to clone proportionally to their antigenic affinity  $f_{i,j}$ , generating a set C of clones
- The set C is submitted to a directed affinity maturation process (guided mutation) generating a mutated set  $C^*$ , where each antibody k from  $C^*$  will suffer a mutation with a rate inversely proportional to the antigenic affinity  $f_{i,j}$  of its parent antibody: the higher the affinity, the smaller the mutation rate:
- Determine the affinity  $d_{k,j} = 1/D_{k,j}$  among  $Ag_j$  and all the elements of  $C^*$ :
- From  $C^*$ , re-select  $\xi$  % of the antibodies with highest  $d_{k,j}$  and put them into a matrix  $M_j$  of clonal memory;
- Apoptosis: eliminate all the memory clones from  $M_j$  whose affinity  $D_{k,j} > d$ :
- Determine the affinity  $s_{i,k}$  among the memory clones:
- Clonal suppression: eliminate those memory clones whose  $s_{i,k} < s$ :

- Concatenate the total antibody memory matrix with the resultant clonal memory  $M_j$
- Determine the affinity among all the memory antibodies from  $Ab_m$ :
- Network suppression: eliminate all the antibodies such that  $s_{i,k} < s$ :
- Build the total antibody matrix  $Ab \leftarrow [Ab(m); Ab(d)]$

Where:

**Ab** - available antibody repertoire ( $Ab \in SN \times L, Ab = Ab(d) \cup Ab(m)$ );

**Ab(m)** - total memory antibody repertoire ( $Ab(m) \in Sm \times L, m \leq N$ );

**Ab(d)** - d new antibodies to be inserted in Ab ( $Ab(d) \in Sd \times L$ );

**Ag** - population of antigens ( $Ag \in Sm \times L$ );

$f_j$  - vector containing the affinity of all the antibodies  $Ab_i$  ( $i = 1, \dots, N$ ) with relation to antigen  $Ag_j$ . The affinity is inversely proportional to the Ag-Ab distance;

**S** - similarity matrix between each pair  $Ab_i - Ab_j$ , with elements  $s_{i,j}$  ( $i, j = 1, \dots, N$ );

**C** - population of  $N_c$  clones generated from Ab ( $C \in S_{NL}$ );

$C^*$  - population C after the affinity maturation process;

$d_j$  - vector containing the affinity between every element from the set  $C^*$  with  $Ag_j$ ;

$\xi$  - percentage of the mature antibodies to be selected;

$M_j$  - memory clone for antigen  $Ag_j$  (remaining from the process of clonal suppression);

$M_j$  - resultant clonal memory for antigen  $Ag_j$ ;

**d** - natural death threshold;

**s** - suppression threshold.

Artificial Immune System response algorithm:

```

NO_Ag = length(Ag);
Ab = Ag;
for i = NO_Ag : N_SegmentationWindowWidth - 1
    LenAg = length(Ab);
    [M, vbD, Dn ] = Answer_Net( M, Ab, i, ds );
    Ab=[Ab, M(1,i+1)];
end

```

The presented algorithm requires some explanations. There are important operations that are responsible for suppressive steps. The steps are called clonal suppression and network suppression, respectively. As far as a different clone is generated to each antigenic pattern presented, a clonal suppression is necessary to eliminate intra-clonal self-recognizing antibodies, while a network suppression is required to search for similarities between different sets of clones. After the learning phase, the network antibodies represent internal images of the antigens.

### 3 Example - Prediction of Blood Gasometry Parameters

In the given example use of the artificial immune network is presented in the prediction of the  $pH$ ,  $pCO_2$ ,  $pO_2$  and  $HCO_3$  parameters. Using a database, functioning for the few years on the Infant Intensive Care Ward of the Polish - American Institute of Pediatrics Collegium Medicum Jagiellonian University of Cracow, an artificial immune network was created, which task is prediction of the arterial blood gasometry parameters ( $pH$ ,  $pCO_2$ ,  $pO_2$ ,  $HCO_3$ ). In the process of training previous values of gasometry, respirator settings and surfactant administration were used as an input data.

Training process of the artificial immune network consists of two phases. First phase is a learning of the artificial immune network. This phase proceeds accordingly to the algorithm presented previously.

Training set is comprised of blood gasometry parameters time series starting from the time  $t$ . Time series are segmented as shown in Figure 1.

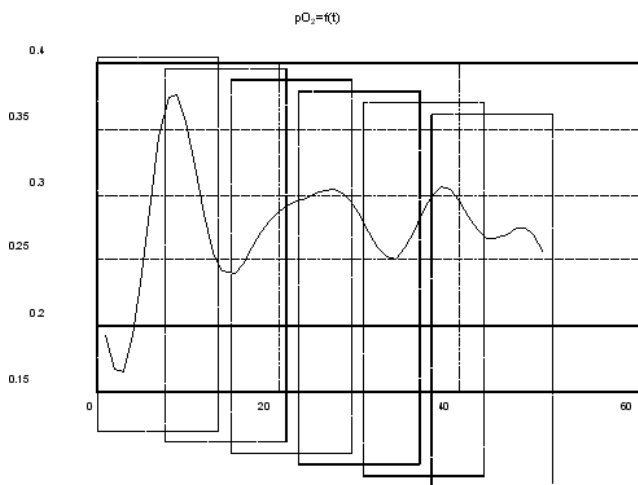


Fig. 1.  $pO_2=f(t)$  time series segmentation process

Second phase tests the network’s generalization abilities by presenting it input vectors from the test dataset.

#### 3.1 Training Process

Dataset for the training process consisted of 480 samples, and parameters of the artificial immune network were set to the following values:

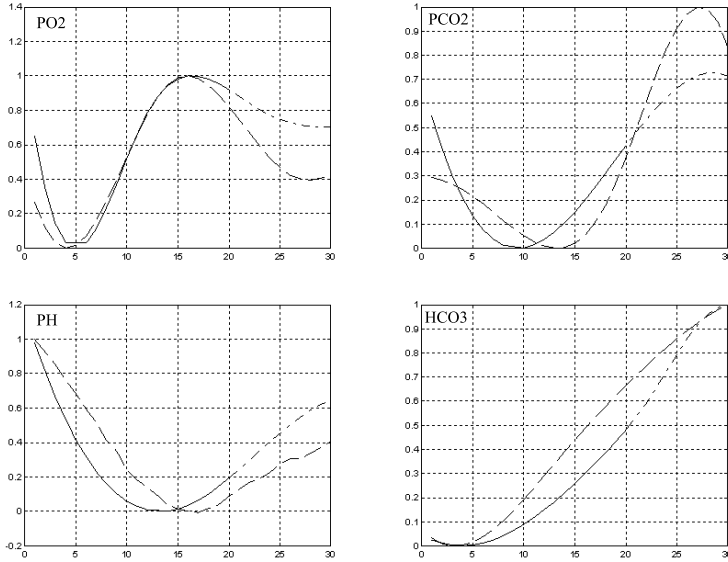
$$n = 4, \xi = 20\%, \sigma d = 0.8, \sigma s = 0.01$$

Stopping criterions of the training process were set to one hundred generations  $N_{gen} = 100$ . One of the more important parameters, which has a major influence on the immune network structure is a ( $\sigma s$ ) parameter. Changing value of the

suppression threshold ( $\sigma_s$ ) influences the generalization abilities of the artificial immune network.

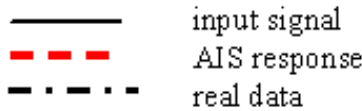
### 3.2 Testing Process

Figure 2 presents example response of the network, for the time series, chosen for the infants blood gasometry parameters data set.



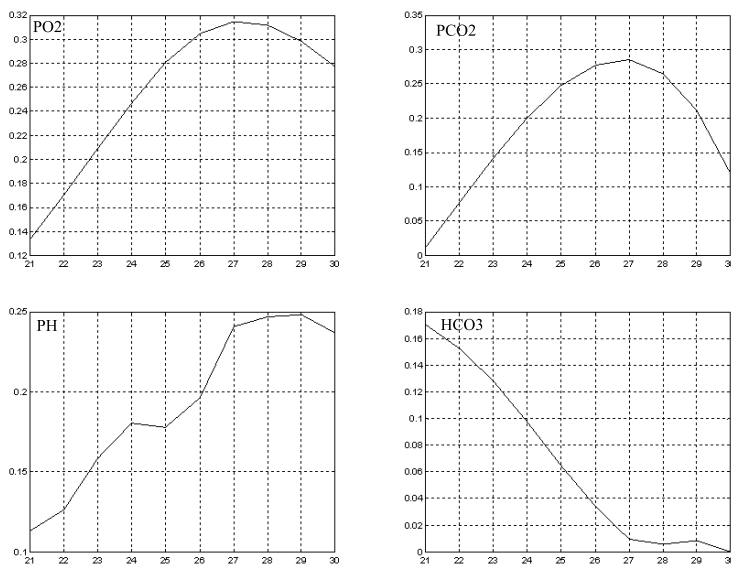
**Fig. 2.** Time series prediction result of  $pO_2$ ,  $pCO_2$ ,  $pH$  and  $HCO_3$  parameters for one of the patients

Legend:



## 4 Summary

In the paper application of artificial immune network in time series prediction of medical data was presented. Prediction mechanism used in the work is basing on the paradigm stating that in the immune system during the response there exist not only antigene - antibody connections but also antigene - antigene connections, which role is control of antibodies activity. The mechanism turned out to be an interesting technique useful in prediction of the time series. Future work aims at increasing systems reliability by improving the algorithms used and acquiring further data of the infants hospitalization.



**Fig. 3.** Prediction error calculated for  $pO_2$ ,  $pCO_2$ ,  $pH$  and  $HCO_3$  parameters

## Acknowledgements

This work was partially supported by the Polish Government through a Grant No. 3T11E04227 from the State Committee of Scientific Research (KBN).

## References

1. Kruczek, P. Assesment of neural networks methods usefulness in prediction of premature neonates respiration insufficiency, Doctoral Dissertation, Collegium Medicum, Jagiellonian University in Cracow, Cracow 2001.
2. De Castro, L. N., Von Zuben, F. J. (2000a) An Evolutionary Immune Network for Data Clustering, Proc. of the IEEE SBRN, pp. 84-89.
3. De Castro, L. N., Von Zuben, F. J. (2000b) The Clonal Selection Algorithm with Engineering Applications, GECCO'00 - Workshop Proceedings, pp. 36-37.
4. Wajs, W., Wais, P., Swiecicki, M., Wojtowicz, H. Artificial Immune System for Medical Data Classification, Proc. of the International Conference on Computational Science 2005, Springer LNCS 3516, pp. 810-812, 2005.