12th IMEKO TC4 International Symposium
Electrical Measurements and Instrumentation
September 25–27, 2002, Zagreb, Croatia

ACCURACY INCREASE OF MULTI-SENSOR MEASURING SYSTEM USING SENSORS DATA FUSION

P. Girão[1], O. Postolache[1,2], M. Pereira[2], H. Ramos[1]

[1]IT/IST, Av. Rovisco Pais, 1049-001 Lisbon, Portugal
[2]ESTS/IPS, Rua do Vale de Chaves, Estefanilha, 2914-508 Setúbal, Portugal

Abstract – The increase of multi-sensor measuring systems' accuracy by fusion of sensors data is analysed. Methods for data fusion and the results they produce are presented and discussed in the context of one application related with environmental monitoring.

Keywords: data fusion; sensors.

1. INTRODUCTION

Data fusion can be defined as the process of combining data or information to estimate or predict entity states [1].

The information integration underlying data fusion can be used for several different purposes. Initially associated with military applications [2-4], data fusion has increasingly been used in other contexts. Many metrological applications of neural networks are precisely in the context of data fusion [5, 6].

The situation addressed in the paper, which is related with the increase of sensor's selectivity and is in the domain of application is environmental monitoring, can be summarised as follows: n quantities are to be identified and measured. For that purpose, a set of n sensors whose outputs are of the same type (e.g. a voltage level) is used. Each sensor informs on a quantity, but due to limited selectivity, their outputs are correlated. How to extract the value of each quantity with increased accuracy? As we will show, it is possible to increase that accuracy by using data fusion methods. The efficiency of the possible solutions is of special interest since the same methodologies can be used in formally identical problems that may arise in other contexts.

2. PROBLEM SPECIFICATION

The work now reported, and briefly introduced in the previous paragraph, was motivated by the need of using low cost, poor selective sensors (of the ion selective electrode (ISE) type) for the identification and concentration measurement of heavy metal ions in river waters. The correlation between sensors output, due to their limited selectivity, is exemplified in Figs. 1 and 2 for two of those sensors. Fig. 1 shows the behaviour of a lead selective electrode, ELIT8231, in the presence of interference ions of cadmium (Cd^{2+}). The sensor is sensitive to Cd^{2+}, its response to these ions being expressed by a catalogue selectivity coefficient $k_{Cd}^{2+}=0,1$. Fig. 2 is a similar representation for a

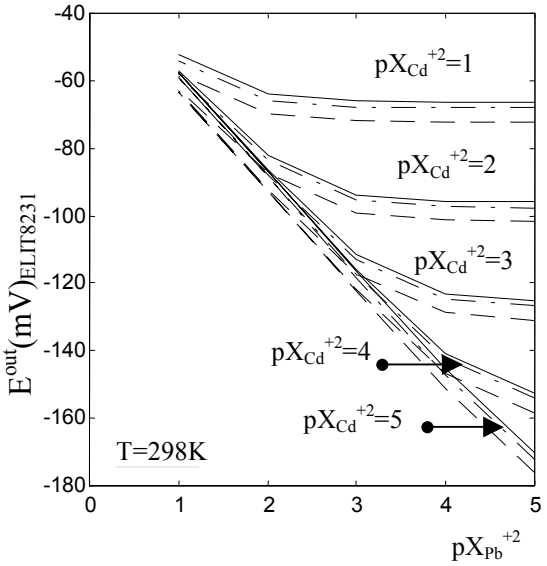

Fig. 1. The evolution of the ELIT8231 response for different concentration values of the Cd^{+2} ions and for different values of the activity coefficient γ: solid curves γ=0,982, dash dot curves γ=0,864, dotted curves γ=0,628

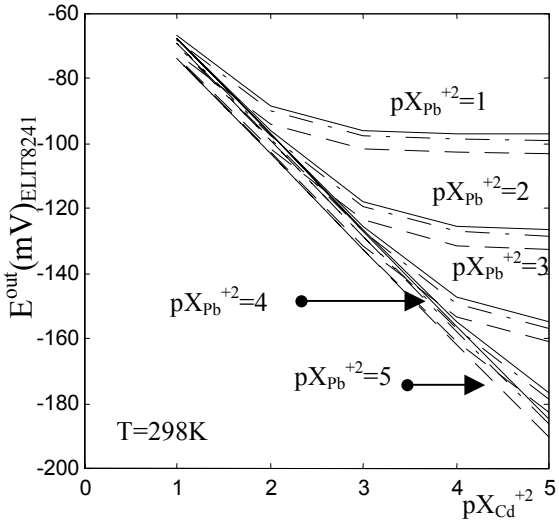

Fig. 2. The evolution of the ELIT8241 response for different concentration values of the Pb^{2+} ions and for different values of the activity coefficient γ: solid curves γ=0,982, dash dot curves γ=0,864, dotted curves γ=0,628

cadmium ion selective electrode, ELIT8241. The selectivity coefficient of this sensor to lead ions (Pb^{2+}) is k_{Pb}^{2+}=0,2. It is clear that to obtain the concentration of each metal with increased accuracy using a set of such sensors, some processing on sensors output must be performed. Two approaches can be followed: (1) to process the output voltage of each sensor and then use the transfer characteristic of each sensor to obtain the correspondent concentration value; (2) to develop a processing scheme that corrects directly the concentration values. The first type of approach can be implemented using either a probabilistic model (e.g. Bayesian, Dempster-Shafer) or weighted decision methods (voting techniques); for the second type of solution, both weighted decision and artificial intelligence methods (neural networks, fuzzy) are adequate. In this paper we will deal only with weighted decision and neural networks based methods. The implementation of the methods is exemplified using the two sensors already mentioned.

2.1. Voting

Considering that the electrode potentials (E_{Cd}^{2+}, E_{Pb}^{2+}) dependence with the activity of the measured ions (a_{Cd}^{2+}, a_{Pb}^{2+}) is expressed by the Nikolsky-Eisenman equation (potentials in milivolt)

$$E_{Pb^{+2}} = E_{ref} + \frac{2.3 \cdot R \cdot T}{Z \cdot F} \cdot \log\left(a_{Pb^{+2}} + k_{Pb,Cd} \cdot a_{Cd^{+2}}\right)$$
$$E_{Cd^{+2}} = E_{ref} + \frac{2.3 \cdot R \cdot T}{Z \cdot F} \cdot \log\left(a_{Cd^{2+}} + k_{Cd,Pb} \cdot a_{Pb^{2+}}\right) \quad (1)$$

(E_{ref} - reference electrode potential; $k_{Pb,Cd}$, $k_{Cd,Pb}$ - potentiometric selectivity coefficients for the $Pb2^{+}$ and Cd^{2+} ion selective electrodes; R – gas constant (8314 J/K); T - absolute temperature; Z - charge on the ion; F - Faraday constant (96500 coulombs/mol)), a *'voting'* type technique is applied in order to produce a better estimation of the concentration of the Cd^{2+} and Pb^{2+} ions. As it is presented in equation (1), temperature can affect seriously the determination of ion concentration. Thus, an additional temperature sensor based on ON401 thermistor is introduced in the system. The voting technique structure is presented in Fig. 3.

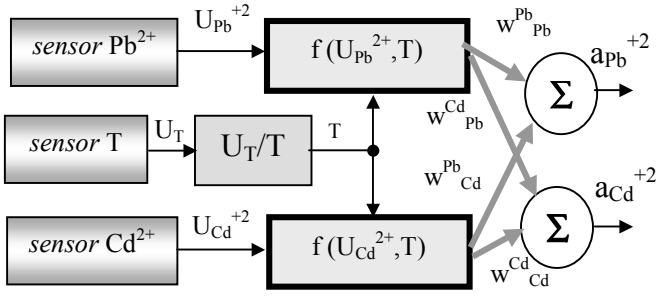

Fig. 3. The *voting* technique structure

According to the structure of Fig. 3, the acquired voltages ($U_{Pb}^{+2}=E_{Pb}^{+2}-E_{ref}$, $U_{Cd}^{+2}=E_{Cd}^{+2}-E_{ref}$) associated to the ISE independent measurement channels are applied to processing blocks whose transfer functions *f* are defined by:

$$f(U_X, T) = 10^{\frac{U_X \cdot F \cdot Z}{2.3 \cdot R \cdot T}} \quad (2)$$

The output values of the Pb^{2+} and Cd^{2+} processing blocks are weighted with the values (w^{Pb}_{Pb}, w^{Cd}_{Pb}, w^{Pb}_{Cd} w^{Cd}_{Cd}) obtained after the calculation of the inverse of the potentiometric selectivity coefficient matrix, $\mathbf{K_{Pb,Cd}}$ defined as:

$$\mathbf{K_{Pb,Cd}} = \begin{bmatrix} 1 & k_{Pb,Cd} \\ k_{Pb,Cd} & 1 \end{bmatrix} \quad (3)$$

$$\mathbf{W_{Pb,Cd}} = \begin{bmatrix} w^{Pb}_{Pb} & w^{Pb}_{Cd} \\ w^{Pb}_{Pb} & w^{Pb}_{Pb} \end{bmatrix} = \mathbf{K}^{-1}_{\mathbf{Pb,Cd}} \quad (4)$$

Based on the above relations the activities of the monitored ions in water solution are obtained.

$$a_{Pb^{2+}} = w^{Pb}_{Pb} \cdot f(U_{Pb^{2+}}, T) + w^{Cd}_{Pb} \cdot f(U_{Cd^{2+}}, T)$$
$$a_{Cd^{2+}} = w^{Pb}_{Cd} \cdot f(U_{Pb^{2+}}, T) + w^{Cd}_{Cd} \cdot f(U_{Cd^{2+}}, T) \quad (5)$$

From (5) it becomes clear that this method requires access to the values of potentiometric selective coefficients ($k_{Pb,Cd}$ $k_{Cd,Pb}$). In the present case the potentiometric selectivity coefficients were determined using the Fixed Primary ion Method (FPM) following the IUPAC's recommendation [7, 8]. Thus, the U_X (e.g. U_{Cd}^{+2}) of cell comprising the ISE and the reference electrode (ELIT003) is measured for solutions of constant activity of the primary ion (e.g. a_{Cd}^{2+}) and varying the activity of the interfering ion (e.g. a_{Pb}^{2+}). The obtained voltage values (U_{Cd}^{+2}) are plotted vs. the logarithm of the activity of the interfering ion (Fig.4). The intersection (A) of the extrapolated linear portions of this plot are used to obtain the value of interfering ion (a'_{Pb}^{+2}) and then the potentiometric selectivity based on relation (6).

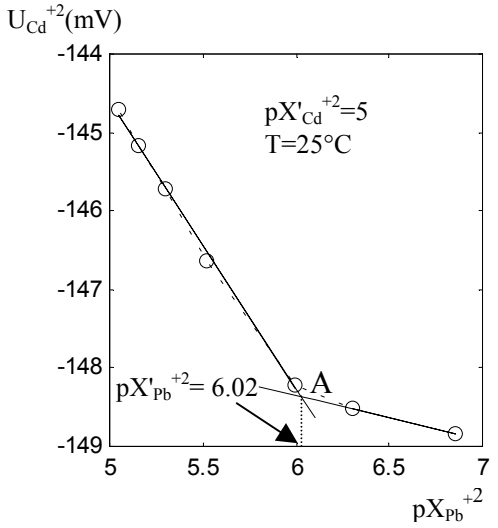

Fig. 4. The evolution of the delivered Cd^{+2} ISE voltage vs. pX_{Pb}^{+2} , where pX_{Pb}^{+2}=-log (a_{Pb}^{+2})

Using IUPAC's specified graphical method the activity value ($a'_{Pb}{}^{+2}$) used to calculate the potentiometric selective coefficient at T=298K, $k_{Cd,Pb}{}^{298}$ is obtained:

$$k_{Cd,Pb}{}^{T} = \frac{a'_{Cd}{}^{+2}}{\left(a'_{Pb}{}^{+2}\right)^{z_{Cd}{}^{+2}/z_{Pb}{}^{+2}}} \quad (6)$$

where:

$$z_{Cd}{}^{+2} = z_{Pb}{}^{+2} = 2$$

$$a'_{Cd}{}^{+2} = 10^{-pX'_{Cd}{}^{+2}}$$

$$a'_{Pb}{}^{+2} = 10^{-pX'_{Pb}{}^{+2}}$$

Using equation (6) and the experimental above mentioned values for ion activities ($a'_{Cd}{}^{+2}$=10ppm, a'_{Pb}=0.95ppm) the obtained $k_{Cd,,Pb}$ of Cd^{+2} ISE is $k_{Cd,Pb}$=0,094≈2. The same procedure is applied to determine $k_{Pb,Cd}$. For T=298K the obtained value is $k_{Pb,Cd}$=0,2. Based on the potentiometric selectivity coefficients are obtained the weighs used to calculate the Cd^{+2} and Pb^{+2} activities and concentration according with the following relations:

$$C_{Pb}{}^{+2} = a_{Pb}{}^{+2}/\gamma_{Pb}{}^{+2} \quad C_{Cd}{}^{+2} = a_{Cd}{}^{+2}/\gamma_{Cd}{}^{+2} \quad (7)$$

where $\gamma_{Pb}{}^{+2}$, $\gamma_{Cd}{}^{+2}$ are the activity coefficients calculated using the Debye-Huckel equation [Nico2000] and the values of ionic strength, I, of the calibration solution (Cd^{+2} as $Cd(NO_3)_2$ and Pb^{+2} as $Pb^{+2}(NO_3)$). The numerical values are $\gamma_{Pb}{}^{+2}$=0,869 and $\gamma_{Cd}{}^{+2}$=0,725.

Using the calculated weights for the particular case of $k_{Pb,Cd}$ and $k_{Cd,Pb}$ ($w^{Pb}{}_{Pb}$=1,041, $w^{Cd}{}_{Pb}$=-0,097, $w^{Pb}{}_{Cd}$=-0,442, $w^{Cd}{}_{Cd}$=1,041) and (5) and (7), the concentrations of Pb^{+2} and Cd^{+2} ions are determined.

The utilisation of the presented technique in river water quality measurements implies on-line compensation of temperature effects on potentiometric selectivity coefficients (ex. $k_{Pb,Cd}$, $k_{Cd,Pb}$) and on activity coefficients.

In the case of potentiometric selectivity coefficients temperature compensation, the $k_{Pb,Cd}(T)$ and $k_{Cd,Pb}(T)$ is determined in laboratory and the matrix $\mathbf{K_{Pb,Cd}}$ is recalculated for the temperature of the tested solution in order to diminish the temperature effects in the ion activity calculation.

Regarding the activity coefficients $\gamma_{Cd}{}^{+2}$ and $\gamma_{Pb}{}^{+2}$, additional information of ionic strength (I) is necessary in order to calculate their values. I is defined by the following relation:

$$I = \frac{1}{2}\cdot\sum_i \left(C_i \cdot z_i^2\right) \quad (8)$$

where C_i is the 'i' ion concentration, z is the 'i' ion valency,

In practical applications the ionic strength of a solution is expressed as a product of the solution conductivity (σ) and η, a factor experimentally determined. Using I determined as a linear function of conductivity, the activity coefficient γ_I of the considered ion is determined by the Debye-Huckel modified equation,

$$\gamma_i = 10^{-\left(0.51\cdot z_i^2\cdot\sqrt{\eta\cdot\sigma}/1+3.29d\sqrt{\eta\cdot\sigma} - 0.1\cdot z^2\cdot\sigma\right)} \quad (9)$$

where η represent the proportional factor experimentally

To compensate temperature influences on γ_i values, temperature compensation is performed on the conductivity transducer used (ISI OLS50).

As it is above presented, *the voting weighted method* is characterised by higher complexity and limited capabilities in terms of disturbance factors compensation. An attractive solution for ISE data fusion with application in field measurement is expressed by multi-input neural network

2.2. Neural Networks

The learning capability of neural networks is used to overcome the limited selectivity of the sensors and thus to measure metal concentrations with increased accuracy.

Multiple inputs Multilayer Percepton type neural network (MLP-NN) [9] is used (Fig. 5) to obtain Pb^{+2} and Cd^{+2} concentrations. Taking into account the temperature influence on ISE output voltages, neural network inputs are ISE output voltages ($U_{Pb}{}^{+2}$, $U_{Cd}{}^{+2}$) and temperature obtained as the output of a temperature processing block (ProcT).

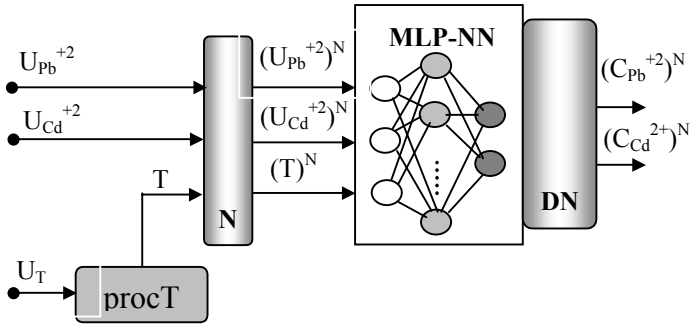

Fig. 5. Fusion architecture based on MLP-NN (Multilayer Percepton Neural Network): N-normalization block, DN – de-normalization block

Referring to the used MLP-NN, the processing elements, called neurons, are distributed in three layers (input layer, hidden layer and output layer) and perform the conversion of normalised voltages and temperature in normalised concentrations of measured ions. The number of hidden neurons is included in 20÷50 interval and the associated function is *tansignoid*. The output layer includes a linear neuron that sums the weighted output values of the hidden neurons. In order to obtain the internal values of the weights and biases that characterise the MLP-NNs a multi-dimensional training set is used. The training set is experimentally determined and corresponds to:

- different concentration of heavy metal ions $C_{Pb}{}^{+2}$, $C_{Cd}{}^{+2}$ = [0,14,1000]ppm (Fig.1, Fig .2)
- different values of activity coefficients $\gamma_{Pb}{}^{+2}$, $\gamma_{Cd}{}^{+2}$ included in the 0,4-0,95 interval obtained for different values of ionic strength I (Fig. 1, Fig. 2)
- different values of temperature, T=[273,298]K (Fig.6).

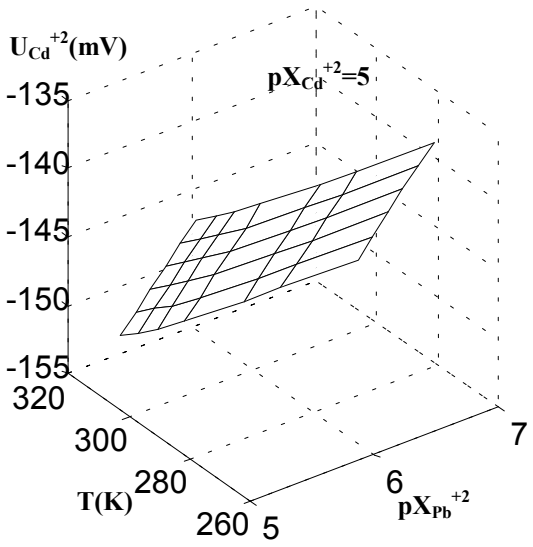

Fig. 6. The evolution of U_{Cd}^{+2} versus temperature for concentration of the primary ion pX_{Cd}^{+2}=5 and concentrations of interfering ion pX_{Pb}^{+2}=5÷7

Based on above-mentioned training set and using the Levenberg Marquardt (LM) training algorithm [NN book], the MLP-NN weights are determined. Considering the problem complexity and also the level of error in the ion concentration measurement using ELIT ISEs (about 5%) [10,11], the training stop condition is expressed by:

$$e_C^{Pb} = \sum_i \left(\left(C_{Pb^{+2}}\right)_i^N \Big|_{MLP-NN} - \left(C_{Pb^{+2}}\right)_i^N \Big|_{real} \right)^2 \le 1E-2$$
$$e_C^{Cd} = \sum_i \left(\left(C_{Cd^{+2}}\right)_i^N \Big|_{MLP-NN} - \left(C_{Cd^{+2}}\right)_i^N \Big|_{real} \right)^2 \le 1E-2 \quad (10)$$

where $\left(C_{Pb^{+2}}\right)_i^N \Big|_{MLP-NN}$ and $\left(C_{Cd^{+2}}\right)_i^N \Big|_{MLP-NN}$ represent the normalised MLP-NN output values of the i^{th} training iteration and the $\left(C_{Pb^{+2}}\right)_i^N \Big|_{real}$ and $\left(C_{Cd^{+2}}\right)_i^N \Big|_{real}$ represent the values of the normalised concentrations as the training target.

After the training phase, weights and biases matrices ($\mathbf{W_1}$, $\mathbf{W_2}$, $\mathbf{B_1}$, $\mathbf{B_2}$) are used for on-line calculation of Pb^{+2} and Cd^{+2} concentrations. Matrices $\mathbf{W_1}$, $\mathbf{W_2}$, $\mathbf{B_1}$, $\mathbf{B_2}$ are processing blocks that are a component of a ion concentration virtual instrument developed in LabVIEW.

Several results regarding the MLP-NN performance in ISE data fusion in order to obtain more accurate values of the ion concentration is presented in the following paragraph.

3. RESULTS AND DISCUSSION

The measurement of ion concentration was performed using a PC based data acquisition system that incorporates the ion selective electrodes, the data acquisition board, a PC and a LabVIEW software that includes the voting fusion block and MLP-NN fusion block.

Considering that Pb^{+2} and Cd^{+2} concentrations in rivers are usually smaller than 10ppm, the above mentioned fusion methods, *voting* and *neural processing,* were tested for a mixed solutions of lead and cadmium ions with concentrations shown in Table I. Sensors' temperature was considered in the 273 K ÷ 298 K interval.

TABLE I. Pb^{+2} and Cd^{+2} for testing solutions

C_{Cd}^{+2}[ppm]	C_{Pb}^{+2}[ppm]
1	0,14; 0,5; 1; 3; 5; 7; 10
3	0,14, 0,5; 1; 3; 5; 7; 10
5	0,14; 0,5; 1; 3; 5; 7; 10
7	0,14; 0,5; 1; 3; 5; 7; 10
10	0,14; 0,5; 1; 3; 5; 7; 10

In both fusion cases the obtained results were compared with the values obtained for direct computation of ion concentration using the Nikolsky--Eisenman relation applied for primary ion ignoring the interference ions.

Fig. 7 presents a comparison between results obtained after Nikolsky—Eisenman direct evaluation and using the voting fusion method for C_{Pb}^{+2}.

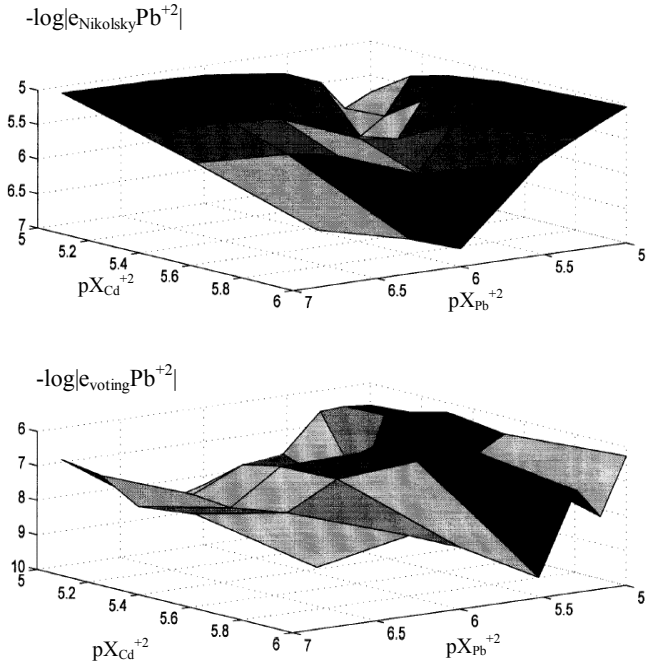

Fig. 7. The evolution of Pb^{+2} concentration errors for the Nikolsky-Eisenman and voting fusion methods (e.g. 5 corresponds to 10 ppm and 7 to 0,1 ppm)

The utilisation of voting method to fuse the information that comes from the two measurements channels (Pb^{+2} and Cd^{+2}) leads to a reduction of ion concentration absolute errors about 16 times in the case of C_{Pb}^{+2} (Fig. 7) and 3 times in the C_{Cd}^{+2} case (Fig. 8) comparing to the situation of the independent ISE measurements (without any information about interfering ions). The graphical

representation of ion concentration errors corresponds to temperature T=296 K and to the ion activity coefficients γ_{Pb}=0,869 and γ_{Cd}=0,956.

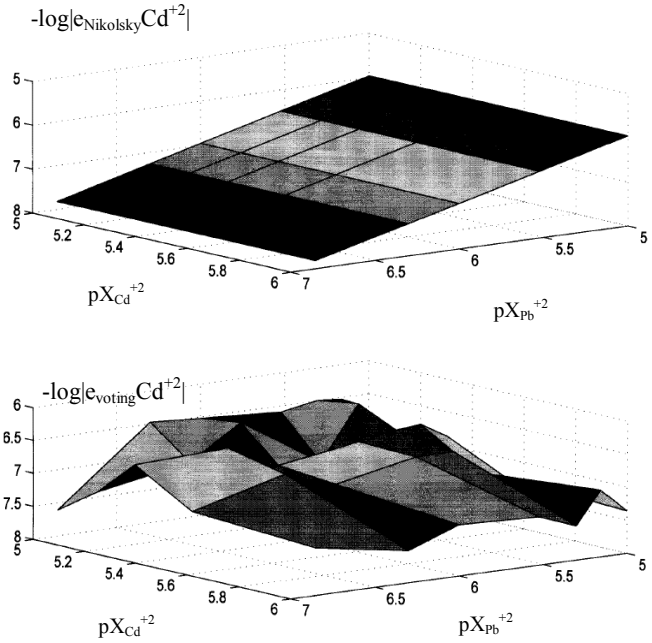

Fig. 8. The evolution of Cd^{+2} concentration errors for the Nikolsky-Eisenman and voting fusion methods

Referring to the temperature variations, the voting fusion method performs the on-line temperature correction through recalculation of the $f(U_{Pb}^{+2}, T)$ and $f(U_{Cd}^{+2},T)$ values but it does not perform any recalculation of the activity coefficient, which implies an increase of ion concentration errors with temperature variations.

Regarding the MLP-NN fusion method, the training set and testing set included the values of ISE channels normalised voltages corresponding to C_{Pb}^{+2} and C_{Cd}^{+2} concentration specified in Table I and different values of temperature T=[273, 298] K. The training set has 3×70 input values and 2×70 target values. The optimum number of hidden neurons obtained after different training is n_{hidden}=35 and the number of epoch necessary to obtain a sum square values less than 0,01 is 720. The graphical representations of ion concentration errors for the MLP-NN fusion solution are presented in Fig. 9 and Fig. 10.

Analysing Figs. 9 and 10, one observes that MLP-NN fusion method errors are less than 0,05 ppm. Comparing with the voting method concentration errors (about 1 ppm) it can be underlined the higher accuracy of the MLP-NN fusion method with application in ion concentration determination.

A second strength of the MLP-NN fusion method is the capacity to operate with sensor data without a previous knowledge about the sensor specific equations that are required in the voting method case.

The more important weakness of the MLP-NN fusion method is the longer calibration time, which includes the time to produce the calibration solutions (different ion concentrations), the ISE data acquisition and the MLP-NN training (weights and biases adjustment).

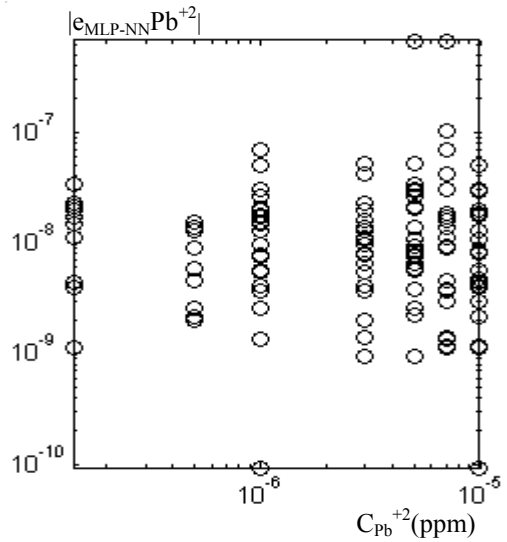

Fig. 9. C_{Pb}^{+2} concentration error associated to the MLP-NN fusion structure (testing phase)

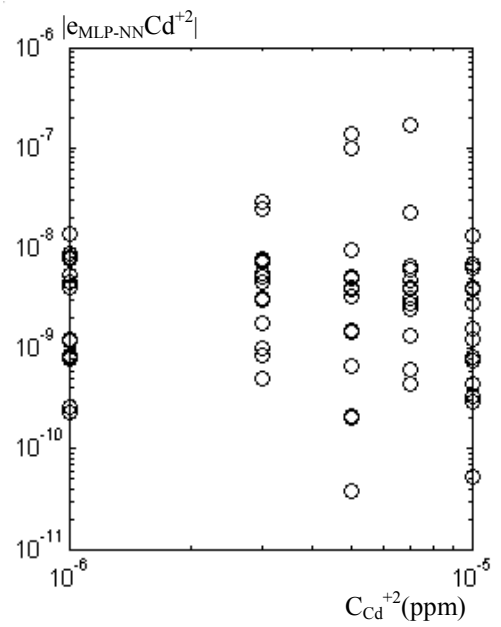

Fig. 10. C_{Cd}^{+2} concentration error associated to the MLP-NN fusion structure (testing phase)

CONCLUSION

Two methods for data fusion have been presented in the context of performance increase of ion selective electrodes ISEs). The electrodes, used here for the measurement of the concentration of heavy metal ions in river water, proved not only to be extremely difficult to handle and operate but also to have a poor selectivity. It was with the purpose of studying how far it was possible to go in ion measurement with ISEs that two processing solutions were considered. The processing involves the use of more than one ISE and

the fusion of the information they convey. While one of the methods uses neural networks, the other can be classified, in our opinion, as a voting type method since it is based on the attribution of weights to the coefficients that relate the concentrations (activities) with electrode's potential.

From the results obtained we conclude that the voting method is a direct method that requires few data but a good knowledge of the sensors to be implemented. On the other hand, neural networks require larger data sets, may not converge but no a priori knowledge of sensors characteristic. In what concerns accuracy, the voting method allows a 10 times error reduction while neural networks are more flexible in error reduction (up to 1000).

ACKNOWLEDGEMENT

This work was supported both by Portuguese Science and Technology Foundation PRAXIS XXI program FCT/BPD/2203/99 and by Project FCT PNAT/1999/EEI/15052. These supports are gratefully acknowledged. We also would like to thank the Centro de Electrotecnia Teórica e Medidas Eléctricas, IST, Lisboa, for its important technical support.

REFERENCES

[1] David L. Hall and James Llinas (Editors), *Handbook of Multisensor Data Fusion*, The Electrical Engineering and Applied Signal Processing Series, CRC Press LLC, Boca Raton (FL), 2001.

[2] R. P. Bonasso, "Expert systems for intelligence fusion", Army conference on Applications of AI to battlefield information management, 1983.

[3] P. L. Bogler, "Shafer-Dempster reasoning with applications to multisensor target identification systems", IEEE Trans. on Systems, Man and Cybernetics, Vol. 17, No 6, pp. 968-977, 1987.

[4] T. J. Hughes, "Sensor fusion in a military avionics environment", Measurement and control, v. 22, n. 7, p. 203, Sept. 1989.

[5] L. I. Perlovsky, M. M. Mcmanus, "Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification", Neural networks: the official journal of the I 1991 v. 4, n. 1, p. 89, 1991.

[6] C. S. Leem, D. A. Dornfeld, S. E. Dreyfus, "A Customized Neural Network for Sensor Fusion in On-Line Monitoring of Cutting Tool Wear", Journal of engineering for industry, v. 117, n. 2, p. 152, May 1995.

[7] Y. Umezawa (Ed.) *Handbook of Ion-Selective Electrodes: Selectivity Coefficients,* CRC Press, Boca Raton, FL, 1990

[8] Y. Umezawa, et. al, "Potentiometric selectivity coeficients of ion-selective electrodes, Part I. Inorganic Cations (Technical Report)", Pure Appl. Chem., Vol.72, No.10, pp. 1852-1856,

[9] S. Haykin, "Neural Networks - A Comprehensive Foundation", Prentice Hall, 1999.

[10]Nico 2000, ELIT ISE Data Sheet, www,nico2000.net