Neural networks application in an AISI 304L intergranular corrosion resistance analysis

D. IACOVIELLO, F. IACOVIELLO and M. MACARIO

***Abstract*—Multi-Layer-Perceptron (MLP) neural network is here considered to predict susceptivity to intergranular corrosion resistance in an AISI 304 L austenitic stainless steel. Six different training algorithms are used. A tempering temperature of 600°C is assumed and 7 tempering duration are investigated (0, 10, 50, 100, 300, 600 and 1000 hours). Intergranular corrosion resistance is investigated by means of electrochemical potentiokinetic reactivation test (EPR) performed between –500 and +200 mV/SCE. Results show an excellent predictive capability of the neural network based on the Levenberg-Marquardt training algorithm considering both activation and reactivation curves.**

***Index Terms*-- Neural network applications, Corrosion testing, Steel.**

I. INTRODUCTION

INTERGRANULAR corrosion in stainless steel is caused by sensitization, a condition of local chromium decreasing: it precipitates in form of carbides (e.g. $Cr_{23}C_6$) along grain boundaries (1).

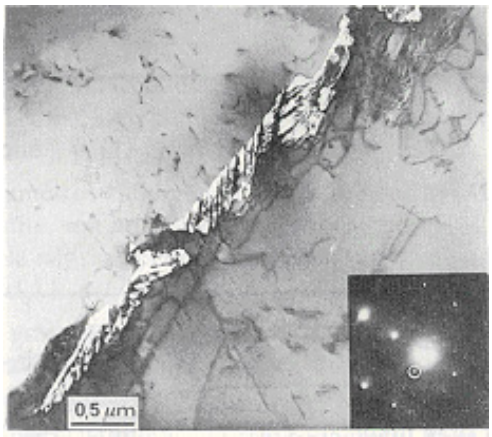

Fig. 1. Carbides $M_{23}C_6$ precipitation at grain boundaries [1].

Since chromium is the essential element in stainless steel corrosion resistance, chromium depleted grain boundaries are susceptible to a local attack, by corrosion that proceeds intergranularly. This phenomenon could be the result of a careless heat treatment, or a welding operation; it is connected to a reheating in a critical temperature range or to a slow cooling from elevated temperatures. Austenitic stainless steels, critical temperatures range between about 500 and 800 °C. This range depends on chemical composition, as carbon or chromium content, with the lowest incubation time that corresponds to about 600-650°C. Degree of sensitization measurements are performed by both destructive and time consuming methods (e.g. ASTM A262 [2]), and non destructive, fast, quantitative, but partially standardized or not standardized electrochemical methods. Among partially standardized methods, the Electrochemical Potentiokinetic Reactivation methods (EPR) [3] are often used to investigate the degree of sensitization, allowing a quantitative analysis of this susceptibility (2).

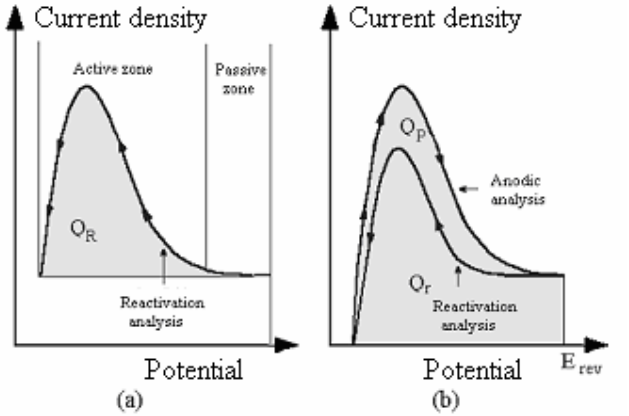

Fig. 2. Schematic polarization curves in single loop (a) and double loop (b) potentiodynamic reactivation.

EPR methods render the functional dependence of current density in response to changes of applied potential [4]. Many types of EPR methods are available. Among them we can consider Single Loop- EPR (SL-EPR) test and Double Loop-EPR (DL-EPR) test. In SL-EPR test the curve is a reverse curve, with the potential scan from positive to negative (2a). In DL-EPR test the curve is polarized anodically at a given rate from the corroding potential to a potential in the passive area. This procedures leads to a passive layer formation on all the surface. Then the scanning direction is reversed and the potential is decreased at the same rate to the corroding

Manuscript received January 17, 2003. This work was supported in part by the MURST (Ministero dell'Università e della Ricerca Scientifica).

D. Iacoviello is with Università di Roma "La Sapienza", Dipartimento di Informatica e Sistemistica "Antonio Ruberti, Via Eudossiana 18, 00184 Roma - Italy (telephone: +39 0644585557, e-mail: iacoviel@dis.uniroma1.it).

F. Iacoviello is with Università di Cassino, Di.M.S.A.T., via G. di Biasio 43, 03043 Cassino (FR) Italy (telephone: +39 0776299681, e-mail: iacoviello@unicas.it).

M. Macario is with Università di Cassino, Di.M.S.A.T., via G. di Biasio 43, 03043 Cassino (FR) Italy (telephone: +39 0776299733,).

potential. This reactivation implies a preferential breakdown of the passive film corresponding to chromium depleted areas, where chromium carbides precipitation takes place (2b). DL-EPR method has the advantage to be substantially independent on the surface finishing.

In this investigation DL-EPR test is conducted on AISI 304L austenitic stainless steel and the degree of sensitization is evaluated as Q_r/Q_p ratio (Area under reactivation peak/Area under activation peak).

Susceptibility to intergranular corrosion depends on the stainless steel chemical composition and its microstructure, on the heat treatment (e.g. tempering temperature and duration (3)) and on the solution physical-chemical conditions (e.g. chemical composition and temperature) [5]. It is quite difficult to characterize the influence of all the variables.

An alternative approach to predicting stainless steel intergranular corrosion resistance is based on the utilization of artificial neural networks. Although the development of mathematical models began in 1960, and they are widely used in many fields such as control system or image analysis and classification [6]-[9], they are not so used in materials science.

In this work we characterize the susceptibility to intergranular corrosion for an austenitic stainless steel AISI 304L tempered at 600°C for different duration by means of DL-EPR test. We try to apply the artificial neural network approach, identifying a training algorithm characterized by a good predictive capability, in order to simulate the sensitization phenomenon in AISI 304L .

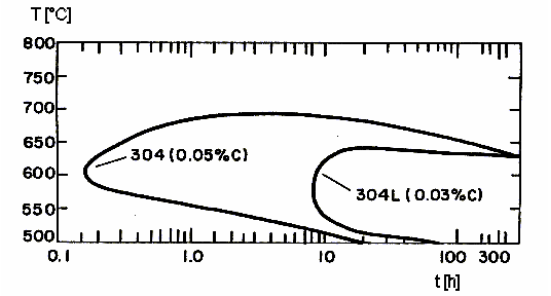

Fig. 3. Austenitic stainless steel AISI 304 TTT (Time-Temperature-Transformation) curves for two different carbon content. The lower the carbon content, the higher the $M_{23}C_6$ carbides incubation time is.

II. THE NEURAL NETWORKS

The neural networks present a structure similar to the one of the human brain. The constitutive parts of a neural network are the input layer x_i, $i = 1,2,...,n$, the hidden layer and the output layer y; to each neurons input a different synaptic weight $w_i, i = 1,2,...n$ is associated. Output values are obtained using the transfer function *f*. Hidden layer neuron output is equal to a weighted sum of input neurons, modified by means of the transfer function *f*. In Fig.4 the outline of the functioning of a neuron is represented [10],[11].

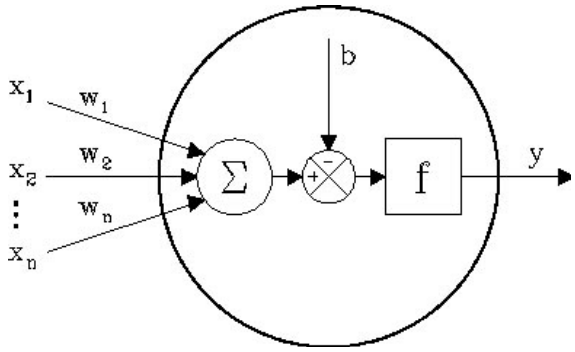

Fig.4. Neuron functioning schematic representation. b represent an activation threshold to which the weighted input is compared; it is usually called *bias.*

For the transfer function f different choices are possible, with the aim of considering a limited function; the most common choices for the transfer function are the step function and the sigmoid; the latter has the advantage of being a derivable function. To use a neural network different phases are needed. In the project phase it is necessary to fix the layers number and the neurons numbers of each layer; in the training phase the synaptic weights are modified to minimize the difference between the real output of the network and the desired one. The generalization phase allows to obtain significant output also with inputs different from the ones used in the training phase.

There are two different kind of training, the supervisioned and the non-supervisioned; the former is based on the availability of couples of input-output divided in two groups: the first one is employed for the training phase, the second for the generalization phase. In the non supervisioned training the network organizes its structure in such way that similar stimuli activate similar neurons as well as far stimuli activate far neurons.

Depending on different applications we can consider various types of neural networks; they differ each other essentially for the structure and the training phase.

The Multi-Layer-Perceptron (MLP) are the most used networks [12],[13]; these networks are characterized by different layers of the feed-forward type, which means that the output of each neurons is joined only to the neurons of the successive layer and that there are no connections of the type feed-back (these are the connection between a neuron and the one of a previous layer); for example, if there are *n* input, *m* output and *q* hidden neurons (5) we will have two weighting matrices W_1 and W_2 of dimensions $n \times q$ and $q \times m$ respectively.

Radial Basis Function (RBF) are networks of the type feed-forward; they have only two layers: the first one is radial basis whose transfer function is a gaussian while the second layer has a liner transfer function. The RBF networks are usually employed in classification problems where the number of possible outputs is not too large.

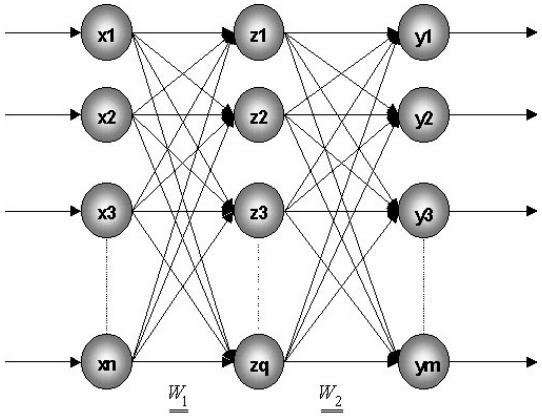

Fig. 5. Example of a Multi-Layer-Perceptron (MLP) network.

Self Organizing networks are able to find the regularities and the correlations between inputs and, according to them, to adapt their future behavior; therefore their training is a non supervisioned one (it is in some way similar to the neuro-biological system). Among the Self Organizing networks, Competitive Network (CN) and Self Organizing Map (SOM) are really common. In the former the transfer function (the competitive function) produces a vector whose components are all null except the *j-th* where *j* is the index of the component with the greater value of the vector processed by the transfer function; the vector of the weights winning the competition is associated to the neuron *j* who is the winner neuron; the vector of the weights associated to this neuron is approached to the input vector: in this way it is increased the probability that with similar input the same neuron will be the winner one. Bias Learning Rule allows, with the offset, to favour the neurons that are rarely winner. SOM networks are similar to the CN ones, but they haven't the offset and update not only the weights of the winning neuron but also the weights of all the neurons that are in a suitable neighbor of the winning one. SOM Network are usually employed for classification problems.

For Learning Vector Quantization (LVQ) Networks the training technique is supervisioned (that is with a training set) but with a competitive structure; there is a competitive layer and a linear one.

Recurrent Networks (RN) allow to connect the neurons in all the possible ways, also with lateral and feed-back connections; the most popular RN are the Elman and the Hopfield Networks. In the former there are three layers , the input, the hidden and the output one who's a feed-back with the input of the hidden layer; the transfer function of the recurrent layer is the sigmoid while the output layer transfer function is linear. The aim of the Hopfield Networks is to simulate a system with some stable equilibrium points; it is difficult to project such a Network which has only the desired equilibrium points and not also spurious ones; the Hopfield Networks do not need the training phase.

MLP Networks are the most popular ones, even if they are not very simple to use because of the presence of various layers useful to the connections feed-forward. The adopted training algorithm for the MLP Networks is the Back Propagation; it is a generalization of the training law of Windrow-Hoff and it is based on the Gradient Discent Algorithm (GDA); let *E* be the mean quadratic error between the real output and the desired one and w_{ij} the generic weight; the GDA can be expressed as follows:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}$$

where η is the learning rate; this means that if $\frac{\partial E}{\partial w_{ij}} > 0$ then $\Delta w_{ij} < 0$; on the contrary if $\frac{\partial E}{\partial w_{ij}} < 0$ then $\Delta w_{ij} > 0$. The network training using the Back Propagation algorithm involves the calculus of the error *E* and the weights updating; this algorithm has some limitations, as the possibility of finding local minimum or long convergence time.

To overcome these problems some variations of this algorithm, based on either euristics or some knowledge of the error surface, are possible. The most common problems regard the weights updating according to this rule:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) + \beta \Delta w_{ij}(t-1)$$

where β is the momentum coefficient; the last term is a kind of memory of the past updating. Another possible variation in the Back Propagation algorithm concerns the possibility of a varying learning rate η; a possible rule can be the following:

$$\eta_{ij}(t) = u\,\eta_{ij}(t-1)$$

where $u \cong 0.7$ or $u \cong 1.05$ according that the two derivatives $\frac{\partial E}{\partial w_{jk}(t)}$ and $\frac{\partial E}{\partial w_{jk}(t-1)}$ have the same sign or not. Montecarlo variation gives a random value to the weights and consider the error *E;* then it gives a random variation to the weights and consider the new error *E'*; if $\Delta E = E' - E < 0$ it accepts the new weights and it gives another random variation to the weights, otherwise it rejects them and consider again the old ones and another random variation. Newton algorithm allows a more rapid convergence than the back propagation one, but it has a greater computational cost because of the Hessian *H*; the updating of the weighting matrix *W* is:

$$W(t+1) = W(t) - H^{-1}(t)g(t)$$

where *g* is the gradient of the error. In order to reduce the computational cost, we consider a Newton algorithm modification (called Levenberg-Marquardt algorithm) where the Hessian *H* is approximated by the Jacobian *J* of the error in weights. Referencing this approximation, the weighting matrix *W* is:

$$W(t+1) = W(t) - \left[J^T J\right] J^T e$$

Finally, another variation of the Back Propagation algorithm is the Resilient Back-Propagation. It is based on the following two rules: the updated weights are increased when

$\frac{\delta E}{\delta w_{jk}(t)}$ has the same sign in two consecutive steps and they are reduced when $\frac{\delta E}{\delta w_{jk}(t)}$ changes its sign in two consecutive steps.

III. Material and Experimental Procedure

A. Material

Wrought austenitic stainless steel AISI 304L is used with the chemical composition given in table I.

TABLE I
AISI 304L Stainless steel chemical composition

C	Cr	Ni	Mn	Si	S	P
0.03	18	10	2,0	1,0	0,03	0,04

Weight percent (wt.%). Fe is balance.

This steel has been investigated both after solubilization and after a tempering heat treatment at 600°C for the following duration:

Tempering time: 10, 50, 100, 300, 600, 1000 hours

These tempering conditions are chosen according to the incubation condition of fig. 2. The worst temperature conditions for AISI 304L correspond to 600°C, and the analysis of the tempering duration influence up to 1000 hours allows us to completely investigate the sensitization phenomenon.

B. Experimental procedures

After the tempering heat treatments, specimens are mounted in an epoxy resin and polished up to 1 μm diamond paste. EPR-DL test are performed in aqueous solution 0.5 M H_2SO_4 + 0.01 M KSCN with the following test parameters:

Temperature: 25°C

Exposed area: 1 cm^2

Sweep rate: 50 mV/min

Potential pause: -450mV/SCE for 2 min

Investigated potential range: -500 - +200 mV/SCE.

In order to control results repeability, DL-EPR test are repeated five times.

As neural networks algorithm, Multi-Layer Perceptron is applied and six different training algorithms performances are investigated:

- Classic Back Propagation;
- Back Propagation with Momentum;
- Variable Learning Rate;
- Resilient Back Propagation;
- Newton algorithm;
- Levenberg-Marquardt algorithm (quasi-Newton algorithm).

The artificial neural network architecture adopted for all the investigated training algorithms is the following:

- Input neurons: 2 (potential and tempering time);
- Hidden neurons: 10
- Output neurons : 1 (current density);
- Training set : 4 x 120 experimental values (after solubilization and after 50, 600 and 1000 hours at 600°C);
- Validation set : 2 x 120 experimental values (after 10 and 300 hours at 600°C);
- Minimum error: 10^{-5}
- Iteration number: 2000
- Transfer function: sigmoid

Simulation has been performed considering the steel after a tempering heat threatment for 100 hours at 600°C.

IV. Results and Discussion

DL-EPR test are characterized by a very high repeability and in fig 6 and 7 voltamperometric curves and Q_r/Q_p values evolution are given.

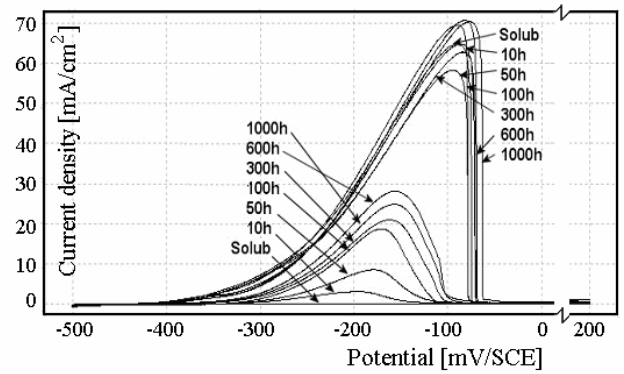

Fig. 6. DL-EPR voltamperometric results for all the seven investigated heat treatment conditions: after solubilization and after 600°C tempering (10, 50, 100, 300, 600, 1000 hours).

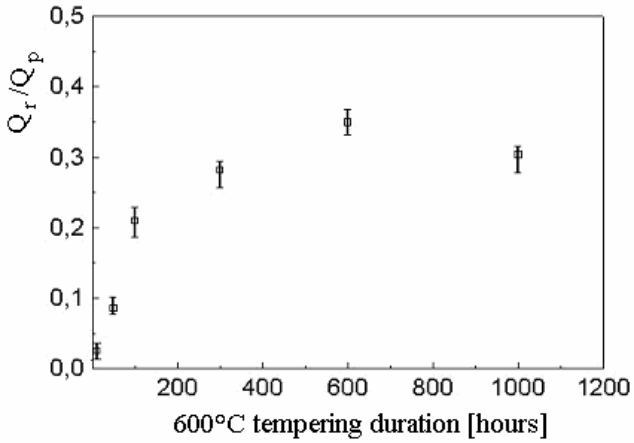

Fig. 7. Q_r/Q_p ratio (Area under reactivation peak/Area under activation peak) evolution with the 600°C tempering duration.

The increasing of tempering duration implies an increasing of chromium carbides precipitation. As a consequence of that, voltamperometric curves are characterized by an increasing importance of reactivation peak with the increasing of the tempering duration. It is confirmed up to a tempering duration of 600 hours. Longer tempering heat treatments at 600°C show an asymptotic Q_r/Q_p ratio, with the chromium carbides precipitation that does not evolve.

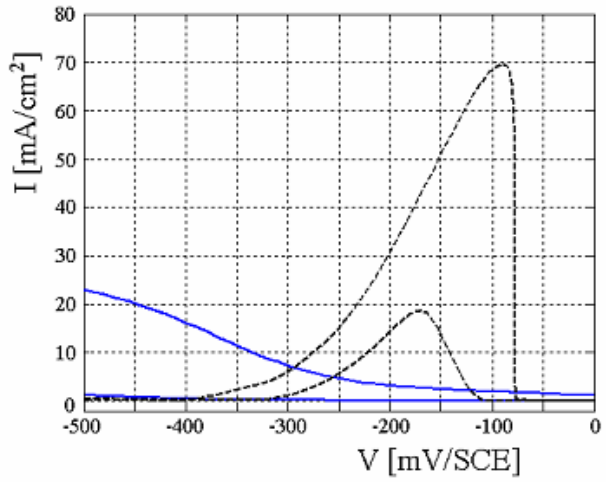

Fig. 8. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Back Propagation training algorithm.

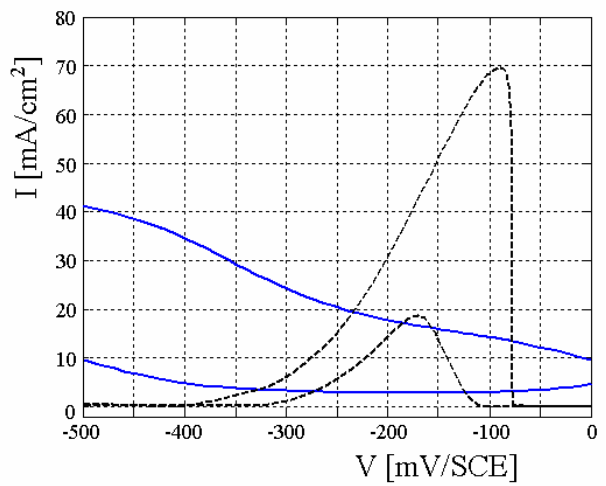

Fig. 9. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Back Propagation with Momentum training algorithm.

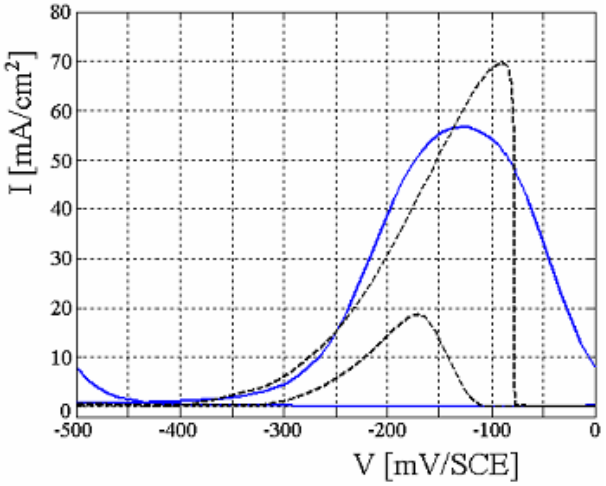

Fig. 10. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Variable Learning Rate training algorithm.

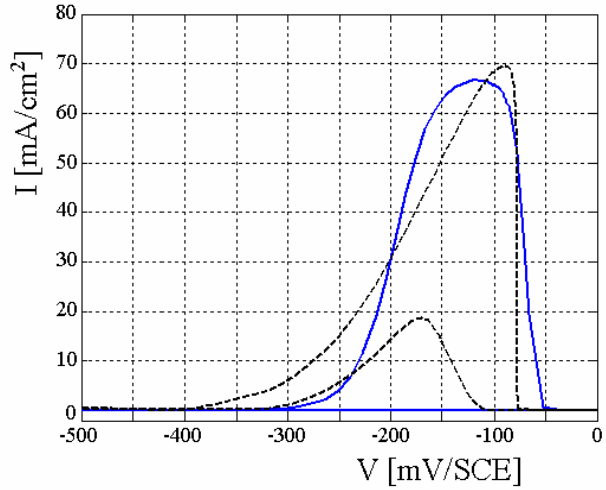

Fig. 11. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Resilient Back propagation training algorithm.

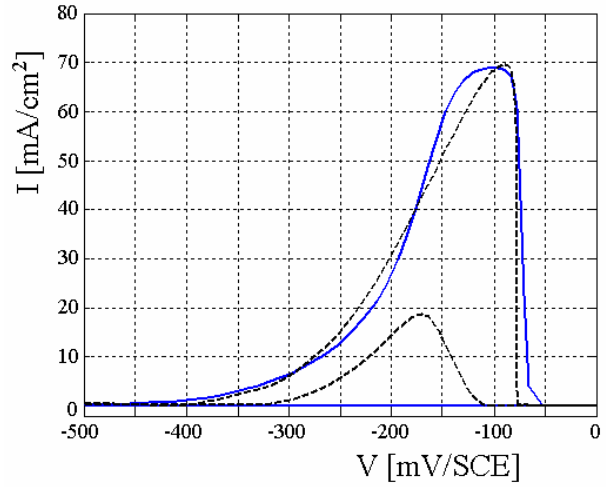

Fig. 12. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Newton training algorithm.

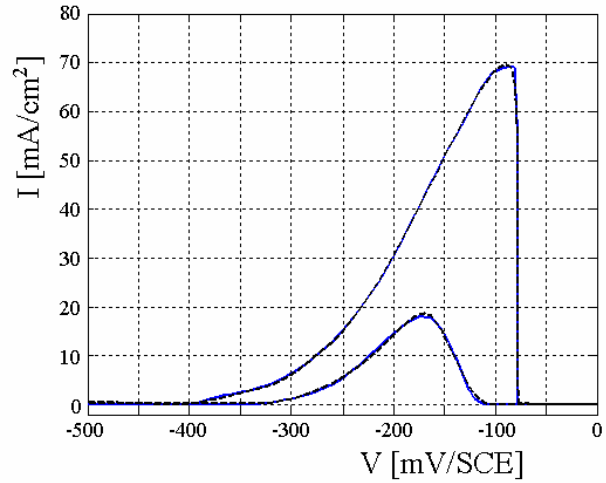

Fig. 13. Experimental (dashed line) and simulated (continuous line) voltamperometric curves for AISI 304L tempered at 600°C for 100 hours. Levenberg-Marquardt algorithm (quasi-Newton) training algorithm.

The six investigated training algorithms are characterized by different generalization capability (figures 8-13). It is possible to observe that both Classic Back Propagation (8) and Back Propagation with Momentum (9) algorithms have a low generalization capability, considering activation curves and reactivation curves. Variable Learning Rate (10), Resilient Back Propagation (11) and Newton (12) algorithms are characterized by a generalization capability that can be considered acceptable for the activation curve, but absolutely not for the reactivation curves. Among them, the Newton training algorithm shows the best activation curve, but the reactivation one is not acceptable.

The Levenberg-Marquardt training algorithm (quasi-Newton algorithm) is characterized by a very good generalization capability (13), considering both the activation and the reactivation curves. The good generalization capability of the Levenberg-Marquardt training algorithm is confirmed when we consider the evolution of Q_r/Q_p ratio as a function of 600°C tempering duration (14). Q_r/Q_p experimental values well correspond with Q_r/Q_p simulated values both for the considered sets (training, validation and simulation set). Differences between experimental and simulated values are lower than the experimental dispersion range (7) and, considering that the training set is quite poor (only 4 elements) and the number of iteration is not high (2000), probably it is possible to improve the generalization capability and obtain simulated activation and reactivation curves that could approximate better experimental curves.

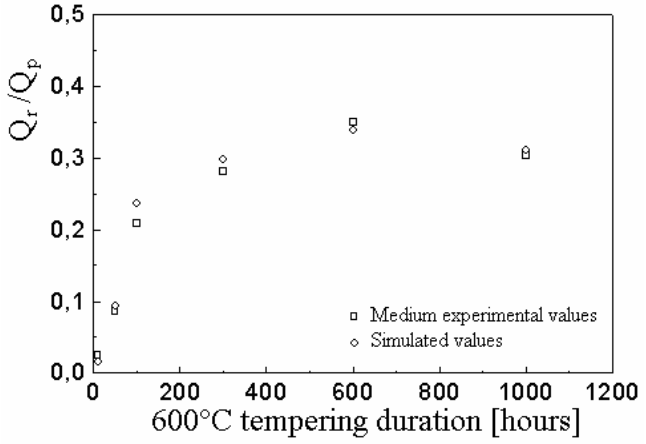

Fig. 14. Q_r/Q_p ratio (Area under reactivation peak/Area under activation peak) evolution with the 600°C tempering duration. Medium experimental values and simulated values obtained using Levenberg-Marquardt (quasi-Newton) training algorithm.

V. Conclusion

DL-EPR is confirmed as a useful method to characterize the intergranular corrosion process in austenitic stainless steels. Austenitic stainless steel AISI 304L is susceptible of intergranular corrosion at 600°C, with the chromium carbides precipitation that develops with a slow kinetics. At the investigated temperature (600°C), tempering duration of 600 hours corresponds to an asymptotic Q_r/Q_p value, with the chromium carbides precipitation that does not evolve more.

In this work the artificial neural network applicability to stainless steel intergranular corrosion analysis performed by means of Double Loop Electrochemical Potentiokinetic Reactivation is investigated. Six different training algorithms are considered, and the Levenberg-Marquardt training algorithm shows the best generalization capability, allowing to obtain simulated activation and reactivation curves that well approximate experimental ones.

Artificial Neural Networks are still not widely used in metals and alloys corrosion analysis. Results obtained in this work confirm that Artificial Neural Networks performances in this field are really interesting and that the increasing possibilities of the generalization capabilities are high.

References

[1] P. Lacombe, B. Baroux, G. Beranger *Les aciers inoxydables*, Les editions de Physique (France) pp. 639, 1990.

[2] ASTM A262-86 *Standard practices for detecting susceptibility to intergranular attack in austenitic stainless steel*, section 3, vol.03.02.

[3] ASTM G 108-92 *Standard test method for electrochemical reactivation (EPR) for detecting sensitization of AISI 304 and 304L stainless steel*, section 3, vol 03.02.

[4] V. Cihal, R. Stefec "On the development of the electrochemical potentiokinetic method", *Electroch. Acta*, 46, pp. 3867-3877, 2001.

[5] S.M. Bruemmer, L.A. Charlot "Development of grain boundary chromium depletion in 304 and 316 type stainless steel", Scr. Metall., 20, pp. 1019-1024, 1986.

[6] T. Parisini, R. Zoppoli, ``Team theory and neural networks for dynamic routing in traffic and communication networks", *Information and* Decision Technologies, 19, 1, pp. 1-18, 1993.

[7] T. Parisini, R. Zoppoli, ``Neural networks for feedback feedforward nonlinear control systems", *IEEE Trans. on Neural Networks*, vol. 5, n. 3,pp. 436-449, 1994

[8] R. Ellis, R. Simpson, P. F. Culverhouse, T. Parisini, ``Committees, collectives and individuals: expert visual classification by neural networks", *Neural Computing and Applications*, vol. 5, pp. 99-105, 1997.

[9] G.L. Foresti, C. Micheloni "Generalized neural trees for pattern classification", *IEEE Transactions on Neural Networks*, 13, 6, pp. 1540-1547, 2002.

[10] T. Khanna, *Foundationsof Neural Networks*, Addison Wesley, 1990

[11] K.J.Hunt, D. Sbarbaro, R. Zbikowski, and P.J. Gawthorp, *Neural Networks for Control Systems*, Automatica, 28, 6, pp. 1083-1113, 1992

[12] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, 1990.

[13] D. Zipser, *A* "Subgrouping Strategy that Reduces Complexity and Speeds up Learning in Recurrent Networks"*, Neural Computing*,1, pp. 552-558, 1989.