

# Prediction the Strain of Traction-aged Polymer Systems from Artificial Neural Networks with Regularization

Hélène Canot, Philippe Durand, Emmanuel Frenod, Bouchra Hassoune-Rhabbour, Valerie Nassiet

**Abstract**—The Liquid Resin Infusion (LRI) is a process that has the greatest development and cost reduction potential for the manufacture of large complex parts which made of composite materials. The viscosity/temperature pair is the essential criterion for the smooth running of the infusion in order to obtain composite parts of quality. However, humidity is a threatening factor for composite materials. Therefore, aging factors and a predictive model of durability were investigated on a new polymer B and second time on A-150, A-185 polymer systems already certified for use in the aircraft and aerospace industry. Tensile tests were carried out at temperatures  $T = -40^{\circ}C, 25^{\circ}C, 70^{\circ}C$ . In this paper, an initial small experimental dataset of 33 samples is used to analyze the strain of polymers systems as a function of aging time, temperature, Young modulus and the breaking stress. In the view of the very small dataset, the strain of polymers systems is predicted by training Levenberg-Marquardt (LM), Bayesian regularization (BR), and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm with a regularized cost function algorithms. The best results have been obtained with the two regularized artificial neural network from very small data set.

**Index Terms**—Artificial Neural Network, Bayesian Regularization, Small dataset, Traction-aged Polymer

## I. INTRODUCTION

Estimating the lifetime of a composite material is a major scientific and technological challenge. Humidity and extreme temperatures are a threat factor for composite materials. This paper is an extended version of [1], we study the aging factor of the composite and we formulate a predictive model of sustainability thanks to artificial neural networks. The purpose of the aging study of polymer systems is to determine what are the irreversible consequences of temperature and water penetration on their chemical structure and on their mechanical properties. Three polymer systems are considered in this study : a new system B, system A polymerized at  $T = 150^{\circ}C$  for two hours with a conversion rate of 89% (noted A-150), system A polymerized at  $T = 185^{\circ}C$  for two hours with a conversion rate of 98% (noted A

185). To meet long-term sustainability criteria, an aging study of these polymer systems is necessary. The type of aging is chosen according to the environmental conditions or the constraints with which the material may be confronted during its commissioning. So water and temperature are two environmental factors that polymer systems are sensitive. Consequently, aging is of the hygrothermal type and the exposure conditions adopted are a temperature of  $T = 70^{\circ}C$ . and a humidity rate of 85%.

In our article we are interested in the characterization of polymer systems aged in traction resulting from the experimental study at the LGP (Laboratoire G nie de Production) of Tarbes. In particular, the data from this study, for polymer B, will serve as a basis for training an Artificial Neural Network (ANN).

In many studies, different modes of aging appear: wet aging by plasticization (Colombini et al. 2002 [8]), by degradation of the polymer by hydrolysis (Ennis et al. 1989 [11]); (Xiao et al. 1998 [35]), by differential swelling linked to concentration gradients (Merdas et al. 2002) [24] but also by damage. The various studies show an influence of hygrothermal aging as a function of the exposure time. Generally, during the hygrothermal aging the mechanical properties of polymers decrease (Dyakonov et al. 1996 [10]); (Popineau et al. 2005 [29]). For short exposure times, a reversible plasticization appears while for long times, swellings and cracks can be identified.

One of the phenomena linked to the penetration of water into the polymer is plasticization, which has consequences on the mechanical and physico-chemical properties of the polymer. The visible changes in mechanical properties are a decrease in elastic modulus and stress accompanied by an increase in elongation at break. Several techniques have contributed to the evaluation of these physico-chemical effects linked to water absorption within polymer systems: the analysis of water diffusion kinetics, rheometry, impact resistance and traction. The experimental study therefore relates to the tensile tests of polymer system B in order to assess the degradation of their mechanical properties over time. In a second time we take these experimental data to establish a predictive model of the behavior of the polymers B aged in traction. This problem is non-linear and an analytical solution is not always easy to obtain. Thus, for reasons of simplicity, advanced identification techniques based on artificial neural networks have been used.

The aim is to develop a neural architecture for predicting the strain of the polymer system through the use of multilayer perceptron (MLP) type with gradient backpropagation. Much

Manuscript received September 20, 2022; revised November 5, 2022  
Helene Canot is associate researcher in Department of Mathematics of Bretagne Sud, Rue Andr  Lwoff 56017 Vannes, France, (email: helene.canot@univ-ubs.fr).

Philippe Durand is assistant professor in Department of Mathematics (mod lisation math matique et num rique), Conservatoire National des Arts et M tiers, 292 rue Saint Martin, 75141 Paris, FRANCE, (e-mail: philippe.durand@lecnam.net).

Emmanuel Frenod is professor at Department of Mathematics of Bretagne Sud, Rue Andr  Lwoff 56017 Vannes, France, (email: emmanuel.frenod@univ-ubs.fr).

Bouchra Hassoune-Rhabbour is assistant professor at ENI de Tarbes, 47 avenue d'Azereix, 65016 Tarbes, France, (email: bouchra.hassoune@enit.fr). Valerie Nassiet is professor at ENI de Tarbes, 47 avenue d'Azereix, 65016 Tarbes, France. (email: valerie.nassiet@enit.fr).

research has proven the usefulness of ANN in modeling, and predicting mechanical properties. Among the many models of ANN, the multilayer perceptron is the most widely applied because it can define the non-linear and complex relationships of the mechanical behavior of composite materials. This type of ANN can produce a relationship between the discrete data and the output values. The non-linear effects of the parameters are automatically captured in the network parameter via the synaptic weights and the threshold.

In recent years, many researchers have used artificial neural networks in the field of composite materials to predict their behavior Zhang (2003) [38] and Goh (1995) [14]. Qingbin et al. (1996) [27] constructed a feedforward neural network with two hidden layers, temperature, effective strain, and strain rate were the inputs, and stress was the output of the neural network. It was able to approximate the constitutive relation for a thermal viscoplastic material. Huber and Tsakmakis (2001) [15] show that the ANN identified physically sets of parameters of the material composite. And it correctly predicted experimentally observed material behavior. ANN make it possible to estimate, at lower digital cost, the level of damage to a composite without resorting to exact calculation.

Mahmoudi (2017) [20] used ANN modeling allowing a good localization and estimation of the damage as well as the prediction of the dynamic response of composite structures totally or locally, damaged while reducing the cost of calculation.

Burgaz et al. (2014) [7] employed ANN method with a feedforward back propagation algorithm for the prediction of thermal stability, crystallinity and thermochemical properties of polyethylene (oxide)/clay nanocomposites. The ANN results confirm that nanocomposites thermal stability increases with the decrease of enthalpy of melting and relative crystallinity.

Doblies et al. (2019) [9] used ANN and Fourier-transform infrared spectroscopy (FTIR)- to predict the mechanical properties, as well as the thermal exposure time and temperature of epoxy resin and composite. It is a novel approach to combine Fourier-transform infrared spectroscopy (FTIR), data processing, and machine-learning (ML) to estimate the material state. The ANN has been trained and has shown the feasibility of predicting the coupled degradation parameters, time and temperature, individually, using only the FTIR spectra.

An ANN was performed by Barbosa et al. (2019) [5] to model the temperature-frequency dependence of dynamic mechanical of thermoplastic polymers of advanced composites. They studied a new Elium<sup>®</sup> acrylic matrix developed by Arkema to evaluate the accelerated test methodology based on time-temperature superposition principle of Carbon Fiber/Elium<sup>®</sup> 150 composites. The learning rule employed by the ANN was Levenberg-Marguardt algorithm, with the gradient descent transfer function into the network. The temperature and frequency dependence were chosen as input parameters and the output parameters provides information about the material properties of the carbon fiber.

Very recently, Adesina et al (2020) [2] examine the potential of ANN for the prediction of mechanical properties, namely density and hardness of graphene nanoplatelet (GNP)/polylactic acid (PLA) nanocomposite developed un-

der various operating conditions of spark plasma sintering (SPS) technique. They employed back-propagation architecture and Levenberg-Marquardt algorithm to predict the mechanical performance in terms of density and hardness property of GNP/PLA nanocomposites.

ANN is inspired by the way the brain processes information. It is defined as a computational model whose design is very schematically similar to the basic operating concept of biological neurons ANN developments have gone through three periods of activity. The first period in the 1940s was due to the work of McCulloch and Pitts (1943) [22]. The second occurred in the 1960s with Rosenblatt's perception theorem of perceptual convergence (1962) [32] and the work of Minsky and Papert (1969) [26] showing the limits of a simple perceptron. Theirs findings have showered the enthusiasm of most researchers, particularly those in the IT community. After a period of silence that lasted almost 20 years, in the early 1980s, the ANNs regained the interest of researchers. This resurgence made it possible to develop the back-propagation learning algorithm for multilayer perceptrons.

The use of a neural network normally requires a large database in order to obtain the best credible model. But the set of experimental data in the field of materials, in particular polymers, is generally limited. In this study, we use an initial small experimental dataset, however ANNs can exhibit problem behaviour in performance with over-fitting and the impossibility of generalization of the neural network. One method for improving network generalization and avoiding over-fitting is to use a technique called regularization. To solve these problems, regularization techniques have been applied. In view of our tiny dataset collected from the polymer B, it has been attempted to provide a strain predicting model based on the Bayesian regularization in combination with Levenberg-Marquardt algorithm and a penalized cost function in combination with BFGS algorithm.

Bayesian regularization provides an interesting performance because Bayesian regularization does not require that a validation data set be separate from the training data set. BFGS algorithm requires more computation in each iteration and more storage than the conjugate gradient methods, but it generally converges in fewer iterations. The BFGS method makes it possible to avoid constructing the Hessian matrix explicitly and to construct instead an approximation of the inverse of the second derivative of the function to be minimized.

Bayesian regularization has been employed to study various problems such as constitutive modeling : in [19] (2003) M. Lefik and B.A. Schreer. proposed a BRANN as a tool for numerical modelling of the constitutive behaviour of a physically non-linear body, or to study magnetic shielding in [17] 2010, where analytical, finite element and BRANN methods was compared to calculate the shielding efficiency of a cylindrical ferromagnetic shield. In [33] Singh et al. (1998) used a Bayesian neural network to predict the yield and tensile strength of rolled steel sheets as a function of chemical composition and processing parameters. Zhang et al. [16] (2002) employed Bayesian regularization neural networks to predict storage and loss moduli of short fiber reinforced composites as a function of material composition

and temperature. Gavard et al. [13] (1996) used BRANN to study formation of austenite during continuous heating of steels. The goal being to predict austenite start and finish temperatures as a function of chemical composition and heating rate. Recently, in Babuska et al. [4] (2016) BRANN predicted fatigue parameters of 75S-T6 aluminum alloys by using strain-life (S-N) curve data. Arzaghi et al. [3] (2017) presents a dynamic risk-based methodology for maintenance scheduling of subsea pipelines subjected to fatigue cracks using Bayesian network inference.

More recently, Pruksawan et al. [31] (2019) proposed the optimization of a very high strength adhesive material based on an active learning model and Bayesian optimization. This combination makes it possible to rely on a small number of experimental data, without having to use data from the literature. They selected the 5 target values closest to experimental reality and then injected them into the initial data set.

Regarding the BFGS algorithm (2017) [6] used six different training artificial neural network algorithms such as f Bayesian regularization, LevenbergâMarquardt and BFGS to predict the failure loads of bonded pultruded composite.

Recently, M. Wiciak-PikuÅa et al. deals in [37] with the phenomenon of tool wear prediction in face milling of aluminum matrix composite materials (AMC), class as hard-to-cut materials. For this purpose, the MLP networks is considered with different activation functions based on cutting force and vibration acceleration measures in the time domain. The BFGS algorithm, which is considered one of the most effective, is selected for training 13 inputs.

Our study has three steps. First, mechanical and physico-chemical characteristic of the polymer B and systems A-150, A-185 were extracted from the thesis work of Laurence Poussines [30]. They are summarized below. Second, tensile tests were performed at three imposed temperatures by the industrialist,  $-40^{\circ}\text{C}$ ,  $25^{\circ}\text{C}$  and  $70^{\circ}\text{C}$ . And thirdly, ANN was carried out to train with the B system and to predict the mechanical properties of the polymer B and systems A-150, A-185 for different aging times.

The aim of our study is to present the prediction of the deformation of polymer as function of temperature, Young modulus, stress and aging time using ANN. On the basis of a small experimental data of polymer system B, we tried to predict, using regularized algorithms, the strain measurements for different aging time.

This paper is organized as follows: section II presents the experimental procedure. Section III concerns the ANN configuration, followed by results and discussion in section IV which concluded the paper.

## II. MATERIAL AND METHODS

### A. Material

In Laurence Poussines work [30], carbon/epoxy composites have been processed by Liquid Resin Infusion (LRI). This process involved selecting a polymer system suited to the parameters imposed by the infusion process. This process makes it possible to impregnate a stack of dry fabrics, without a rigid counter-mold and without autoclaving, only by vacuum pulling. Low viscosity around 100mPa.s and temperature are the essential criteria for the success of the

infusion in order to obtain composite parts of quality. The function that these parts must provide is reliability in the environment because the fuselage structure is subject to temperature variations between  $60^{\circ}\text{C}$  and  $+90^{\circ}\text{C}$  and this temperature range must be ensured even after aging.

Two epoxy/amine systems are used in this work. System A was developed primarily for the injection and reinforcement infusion processes. According to [Kiuna 2002] [18], The system A is a monocomponent. It consists of a stoichiometric blends of a tetrafunctional epoxy prepolymer, the tetraglycidyl methylene dianiline (trademark TGMMA) and two hardeners, The 2,6-diethylaniline and le 2-isopropyl-6-methylaniline. The respective chemical structures of epoxy prepolymer and hardener are given in Scheme 1 and 2.

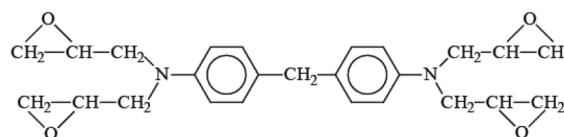


Figure 1. Chemical structure of TGMMA.

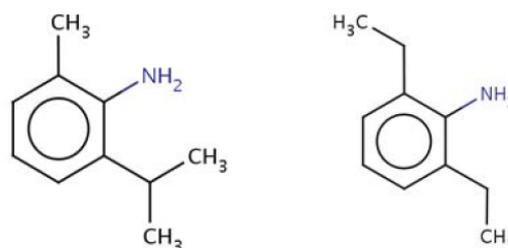


Figure 2. Chemical structure of hardeners present in system A.

The polymer system B is used for the training phase of the ANN and to predict strain of the polymer for different aging time. The system B is composed of a resin and a hardener purchased by Sicomin. Based on supplier data, the resin is a mixture between Bisphenol A DiGlycidylEther (DGEBA) and N, N-Diglycidylaniline, Figure 3

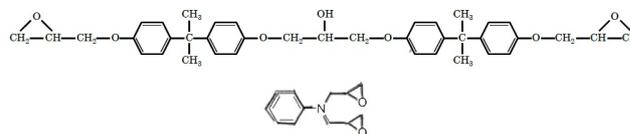


Figure 3. Molecules present in resin B.

Two elements make up the hardener: 4,4-methanedioldicyclohexanamine and 3-(aminomethyl)-3,5,5-trimethylcyclohexanamine, the structures are shown schematically in Figure 4:

### B. Cured polymersâcharacterization before aging

The infusion of system A begins with preheating the resin to  $80^{\circ}\text{C}$ . This then passes through a heating system maintained at  $110^{\circ}\text{C}$  to lower viscosity of system A up to 80 mPa.s then to diffuse in the preformed room maintained

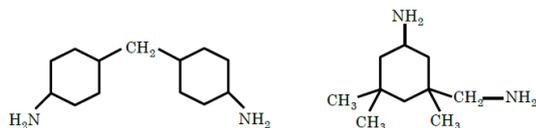


Figure 4. Chemical structure of the substances present in the hardener.

at  $130^{\circ}\text{C}$  in an oven or simply by heating lamps. Finally, the polymerization takes place at  $150^{\circ}\text{C}$  with these same heating systems for 2 hours. Adapting system A [30] to the infusion process generates this polymerization temperature that must not be exceeded. In this case, the polymer system obtained is incompletely crosslinked with a rate of conversion of 89% measured by Infra-Red Spectrometry (IRTF).

An optimized polymerization cycle (2h at  $185^{\circ}\text{C}$  after a rate at  $3^{\circ}\text{C}/\text{min}$ ) made it possible to obtain a rate of high conversion of 98% measured by IRTF as well as a glassy temperature  $T_g$ , determined by Differential Scanning Calorimetry (DSC), of  $211 \pm 3^{\circ}\text{C}$ . These properties are obtained with oxidation on the extreme surface of the samples produced. For the infusion of system B, an optimized polymerization cycle (2h at  $100^{\circ}\text{C}$  + 3h at  $140^{\circ}\text{C}$  at  $1^{\circ}\text{C}/\text{min}$  between isotherm) made it possible to obtain a high conversion rate of 98% measured by IRTF as well as a  $T_g$  of  $135 \pm 3^{\circ}\text{C}$  measured by DSC. These properties are obtained without oxidation on the surface of the samples.

Tensile tests are carried out on  $2 \times 12 \times 45 \text{mm}^3$  rectangular-shaped samples. For each testing temperature and each polymer, 3 test pieces were tested using an INSTRON type machine, equipped with a 5000 N load cell and an INSTRON extensometer, at a speed of  $0.2 \text{mm}/\text{min}$ . The study of elastic modulus and breaking stresses, strain at break by tensile tests at the three temperatures imposed by the manufacturer ( $40^{\circ}\text{C}$ ,  $25^{\circ}\text{C}$ ,  $70^{\circ}\text{C}$ ) showed a worthy difference in behavior between the two polymers, see Table I :

Table I  
MECHANICAL CHARACTERISTICS OF THE TWO SYSTEMS AT DIFFERENT TEMPERATURES.

Polymer system	Temperatures	Young's modulus (Mpa)	Stress (Mpa)
A-185	$T = -40^{\circ}\text{C}$	$3390 \pm 143$	$43 \pm 11$
A-185	$T = 25^{\circ}\text{C}$	$2743 \pm 32$	$69,8 \pm 11$
A-185	$T = 70^{\circ}\text{C}$	$2269 \pm 113$	$59,2 \pm 3$
B	$T = -40^{\circ}\text{C}$	$3643 \pm 127$	$62,7 \pm 22$
B	$T = 25^{\circ}\text{C}$	$2947 \pm 223$	$73,9 \pm 5$
B	$T = 70^{\circ}\text{C}$	$2062 \pm 407$	$49,5 \pm 7$

System B behaves brittle at low temperatures and ductile at high temperatures whereas system A  $185^{\circ}\text{C}$  behaves brittle whatever temperature. Indeed, elongation at break for system A are lower values compared these of system B. The Young's modulus is similar between the two cured epoxy systems and decreases as the temperature rises. The breaking stresses reach a maximum at ambient temperature but see their values decrease at  $40^{\circ}\text{C}$  and  $70^{\circ}\text{C}$ .

### C. Experimental data set

A total of 3 data sets with 33 points used for the development of the neural network model were collected from experimental values of the system B.

Each data point consisted of the following variables measured in the experiment : stress, percent strain and modulus of elasticity for the aging time sequence:  $t = \{0, 24, 48, 168, 336, 720, 2160, 4320, 6480, 8640, 10800\}$  hours and the temperatures of  $-40^{\circ}\text{C}$ ,  $25^{\circ}\text{C}$  and  $70^{\circ}\text{C}$ . The experimental data obtained from tensile tests for the B system studied at different temperatures and different aging times are shown in Figures 5, 6 and 7.

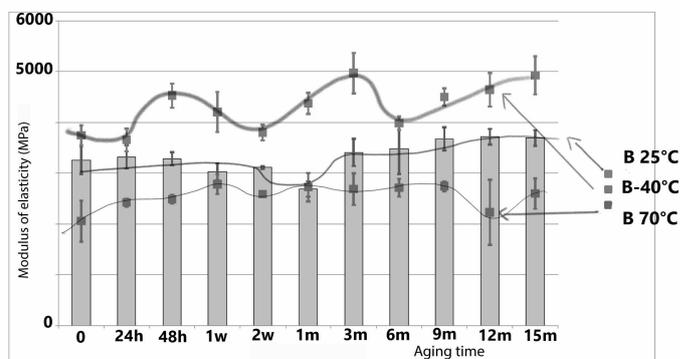


Figure 5. Evolution of the elastic modulus of system B as a function of the aging time and the test temperature.

Tests carried out at room temperature show a slight decrease in Young's modulus up to one month of aging, then, after that, a slight increase in it. Its stress and strain are constant except for uncertainties during the first month and then decrease beyond. The samples undergo from the first stages of aging plasticization due to the penetration of water but which remains relatively low and which leads to a reduction in the Young's modulus. But at longer times, the system becomes rigid and the network is irreparably affected following the hydrolysis in operation. infrared evidence.

Rigid as supported by increase again of elastic modulus and decrease of stress and strain at break.

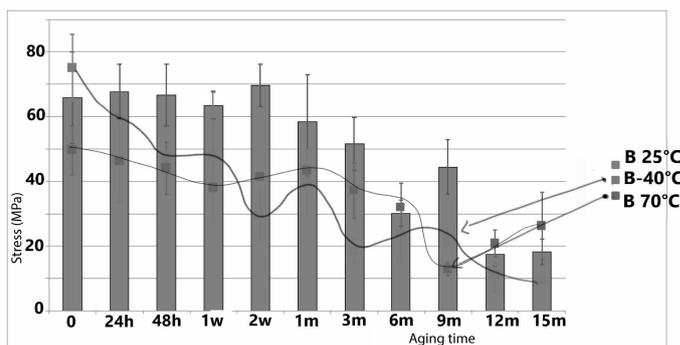


Figure 6. Evolution of the breaking stress of system B as a function of the aging time and the test temperature.

The tests carried out at  $-40^{\circ}\text{C}$  weaken the polymer network even more, which sees its Young's modulus increase and its stress and strain at break decrease. On the other hand, at  $70^{\circ}\text{C}$ , the modulus of elasticity and the tensile stress are lower and the deformation is higher than the values taken at  $25^{\circ}\text{C}$ . The test pieces seem more ductile but a certain brittleness is notable due to the lower values of the

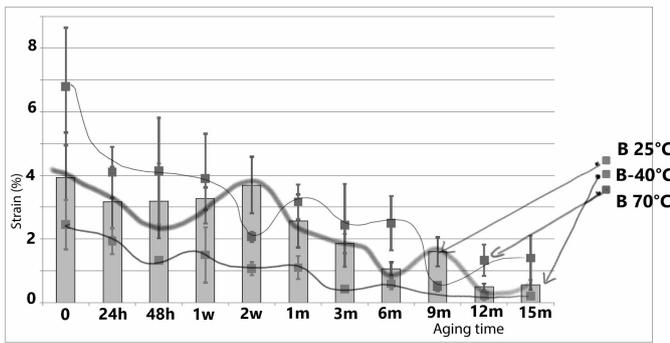


Figure 7. Evolution of the deformation of system B as a function of the aging time and the test temperature.

tensile stress. Let us note all the same an increase in the breaking stress and the strain for aging times greater than 9 months. Indeed, the test temperature is close to the zone of macromolecular mobility. The material is therefore in its ductile zone which generates an increase in stress and strain.

At longer aging times, the network is irreparably affected by oxidation as evidenced by IRTF (support informations) and hydrolysis with chains scissions leading to a mass loss during desorption test in oven at 70°C .

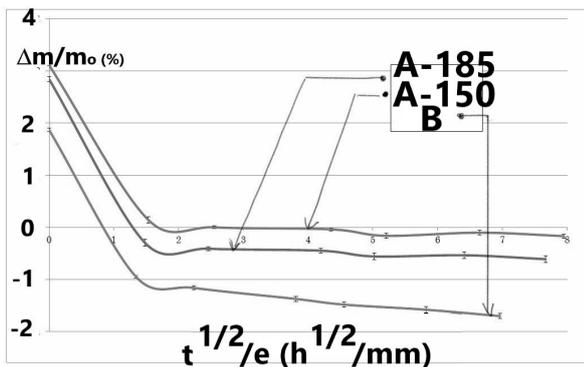


Figure 8. Desorption curve at 140°C for the three systems.

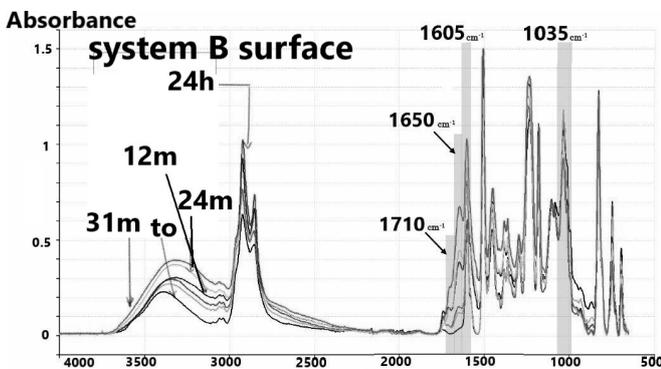


Figure 9. Absorbance spectra of system B at the surface as a function of the number of waves ( $cm^{-1}$ ) and as a function of the aging time.

Also noteworthy is the appearance, increase and decrease of several absorption bands. Several phenomena are at the

origin of these modifications. Infrared analysis shows a decrease in the concentration of CH bonds (between 2700 and 3000  $cm^{-1}$ ) and an increase in the absorption bands of C = O at 1650 and 1600  $cm^{-1}$  characteristic of amides. The chemical reaction leading to the formation of amide under the action of oxygen is shown in Figure

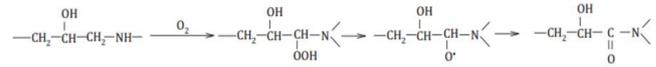


Figure 10. Reaction mechanism of amide formation (Pei 2011) [28].

D. Presentation of the database

In this study, an ANN was implemented with two different algorithms for regularization conditions and compared first to Levenberg-Marquardt algorithm without regularization. As the data set is small we need all the information to feed the neural network. We know that when the tensile tests a certain value of stress is applied on pieces and percent strain is the expected result. A feed-forward network with one layer of hidden neurons was built to represent the complexity the nonlinear nature of the problem.

The input layer consisted of four inputs: the stress, Young’s modulus, aging time and temperatures : thus the total data set of system B consists of 4x33 matrix defining four attributes for 33 different measurements. The targets are the 33 measurements of the corresponding strains of the polymer. The feed-forward neural network model maps the functional relationship between the four parameters and the strain. An optimal number of neurons in hidden layer is selected by testing different choice of neurons number. The data are summarized in the attached table:

Table II  
VARIATION RANGE OF THE POLYMER B.

Time (months)	T °C	E (MPa)	$\sigma$ (MPa)	$\epsilon$ (%)
0 - 15	-40, 25, 70	2000 - 5000	10 - 75	0.1 - 6.9

The model parameters are adjusted during the model calibration phase in order to minimize the error between the value obtained by the network and that normally obtained. In this study, in order to avoid overfitting of the trained ANN, two regularization methods were applied and compared. The first method of regularization consists in modifying the performance function which is the mean sum of squares of the network errors by adding a term that consists of the mean of the sum of squares of the network weights. This method is in combination with BFGS algorithm. The second method consists of Bayesian regularization which determines the optimal regularization parameters in an automated fashion coupled with Levenberg-Marquardt algorithm.

E. Prediction of the strain based on neural network for polymer systems

First, the ANN model was developed and its performance assessed on a dataset of 33 samples of the original dataset. In our case, the database being very small we need as many data as possible for training. Then we divided randomly

the data in two sets, given the small data set, everything is done so that BR trains on the maximum amount of data : a separate training set and a testing set. We take the ratio 90% of the data dedicated to training the network and 10% of the data dedicated to testing the network. Then we compare the performance of the three algorithms BR, BFGS regularized and Levenberg Marquardt. We will see that the two regularized algorithms offer better results than Levenberg Marquardt. After this, for the following we will only use these two regularized algorithms BR and BFGS.

Then, to assess the predictive quality of the neural network, we provide the last measurement of polymer B for the temperature of  $70^{\circ}C$  after an aging time of 15 months and secondly a smaller dataset with 28 samples and strain predictions for aging time  $t = \{3, 6, 9, 12, 15\}$  months.

Having said that, we implement the ANN modeling using Deep Learning Toolbox of MATLAB (R2020a edition), with the Matlab commands : `trainlm`, `trainbr`, `tansig`, `mapminmax` and `trainbfg`. The function "trainbr" that performs Bayesian regularization backpropagation disables validation stops by default because the validation is usually used as a form of regularization, but "trainbr" has its own form of validation built into the algorithm. In other words, the Bayesian error is not just based on how well the model is performing on the dataset, it is also based on how large the weights are. The larger the weights, the higher the error. In fact, during training, validation may never allow the network to explore larger weights, see [21]. "trainbfg" is a network training function that updates weight and bias values according to the BFGS quasi-Newton method and "trainlm" for Levenberg-Marquardt method.

Before training, it is useful to scale the inputs and targets so that they always fall within a specified range. This is necessary to avoid premature saturation of the activation function and allows synaptic coefficients to be kept within relatively small intervals. And it is also about reducing all the inputs of the same order of magnitude which improves the convergence of the algorithm. MATLAB automatically rescaled all input and output variables using the `mapminmax` function such that they resided in the range  $[\hat{a}1, +1]$ . So each variable is normalized in this range using the equation to improve the accuracy and efficiency of calculation:

$$x_n = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (1)$$

where  $x_n$  is the normalized value of the corresponding  $x$ ,  $x_{max}$  and  $x_{min}$  are the maximum and minimum values of  $x$  respectively.

### III. ARTIFICIAL NEURAL NETWORKS CONFIGURATION

#### A. ANN architecture

The most widely used ANN in the community is the multilayer perceptron (MLP), also called feedforward backpropagation. In Figure 11, we see the fully connected network which is divided into layers. In our study, the input layer corresponds in  $p=4$  independent variables and covariates. The input variables are associated with each of  $N$  neurons in a hidden layer by using weights ( $w_{kj}, k = 1, 2, \dots, N$ ) and a

bias specific to each neuron. The number of hidden layer depends the training process. The input vector of independent variables  $p_i = p_1, p_2, p_3, p_4$  is related to the output  $y_i$ .

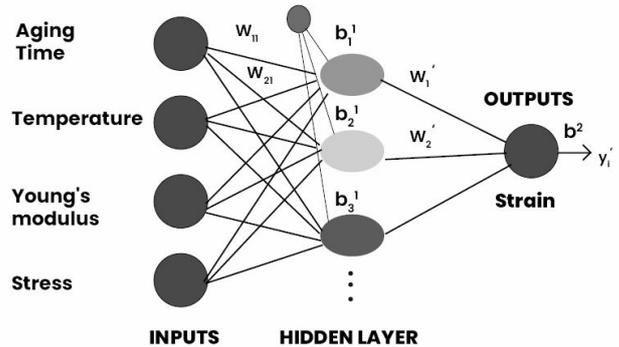


Figure 11. Artificial neural network design with 4 inputs

#### Step 1

For  $N$  neurons in hidden layer of the ANN and appropriate biases:  $b_1^{(1)}, b_2^{(1)} \dots b_N^{(1)}$ , then input values for neuron  $k$  prior to activation is expressed linearly as  $b_k^{(1)} + \sum_{j=1}^4 w_{kj} p_j$ . We applied values to the input in each neuron an activation function which is defined by:

$$f_k(b_k^{(1)} + \sum_{j=1}^4 w_{kj} p_j). \quad (2)$$

#### Step 2

Now, the activated output from the hidden layer is sent to the output layer as  $\sum_{ki=1}^N w'_k f_k(b_k^{(1)} + \sum_{j=1}^4 w_{kj} p_j) + b^{(2)}$  with the weights specific to each neuron  $w_k$  and the bias parameters  $b^{(1)}$  and  $b^{(2)}$  respectively in the hidden and output layers. And at the end, the quantity is activated with the function  $g(\cdot)$  which is  $g[\sum_{ki=1}^N w'_k f_k + b^{(2)}] = a^2 = y'$ , and becomes the predicted value  $y'_i$  of the target variable in the training set as:

$$y'_i = g[\sum_{ki=1}^N w'_k f_k(b_k^{(1)} + \sum_{j=1}^4 w_{kj} p_j) + b^{(2)}], \quad (3)$$

$j = 1, 2, \dots, R \quad k = 1, 2, \dots, N.$

where  $w'_k = \{w_{kj}\}$ .

In our case, we used the sigmoidal activation function such as tangent hyperbolic and logit in the hidden layer. This property shows the interest of neural networks compared to other approximators such as polynomials whose output is a linear function of adjustable parameters: for the same number of inputs, the number of adjustable parameters to be determined is lower for a neural network only for a polynomial. The activation function at the output layer  $g$  is linear.

#### B. Orverfitting and Regularization

As part of machine learning, there are different learning techniques, including supervised learning. We will find a

function  $f$  susceptible, at best according to a criterion to be defined, to reproduce  $y$  having observed  $p$ . The main goals of ANN is to find the predictive algorithm that will provide the best performance from the data available. Performance measurement techniques assess the extent to which predictions obtained through a model approach actual data. Since it is impossible to use new and unperceived data to make a comparison, the available data is divided into training and test subsets. The training subset contains data from which the model is built or trained. We do not focus on the network's ability to be optimal on training data. We want to see the performance of the network on unknown data. This is the whole point of the test subset which is used to assess these performances. There is also a need to improve the model by adjusting its hyperparameters such as the number of hidden layers or the number of neurons. If we make an adjustment of these hyperparameters on the test subset, there is a risk of overfitting the model. If the function learned by the ANN is too finely adjusted to the data, it learns the peculiarities of the training base to the detriment of the underlying model: the neural network is over-adjusted. Over training can be seen as the fact that the model will be more efficient on known data, and much less efficient on new data.

There are two families of methods to prevent overfitting: passive methods and active methods. The philosophies of these two families of methods are different.

- Passive methods try to detect a posteriori overfitting to remove bad models. Among the most classic methods are the use of a validation base during learning, and measures of information criteria.
- Active methods intervene during the training phase to prevent the model from overfitting. Regularization methods such as premature termination or penalization fall within this framework.

Regularization methods are called active because they do not seek to limit the complexity of the network, but they control the value of the weights during learning. It becomes possible to use models with a high number of weight and therefore a complex model, even if the number of learning examples is small. Several regularization methods exist in the literature, in our case, we use active method such as Bayesian Regularization which we compare to another method of regularization which modifies the performance function with BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm.

In the Bayesian approach, all the parameters, in particular the network weights, are considered as random variables from a probability distribution, the weights are assigned a probability fixed a priori, and, once the training data have been observed, this a priori probability is transformed into posterior probability thanks to Bayes' theorem. In the following section we review Bayesian techniques, applying by (MacKay 1992 [23]; Dan. Foresee and Hagan 1997 [12]) to optimize regularization.

1) *Bayesian Regularization*: The training process is carried out by minimizing a function  $F$  named cost function, computing the distance between real and predicted data, this

function determines the objective to be reached. The function writes:

$$F = E_D(D|w, M) = \frac{1}{N} \sum_{i=1}^n (e_i)^2 = \frac{1}{N} \sum_{i=1}^n (y'_i - y_i)^2 \quad (4)$$

where  $E_D$  is the mean sum of squares of the network error,  $D$  is the training data set and  $M$  is the specific functional form of the neural network architecture.

In Bayesian Regularization (BR), an extra term,  $E_w$ , is added by the neural network to the objective function which penalizes large weights in anticipation to reach a better generalization and smoother mapping. A gradient-based optimization algorithm is then applied to minimize the function:

$$F = \beta E_D(D|w, M) + \alpha E_w(w, M) \quad (5)$$

where  $E_w(w, M)$  is the sum of squares of architecture weights,  $M$  is the ANN architecture and  $\alpha$  and  $\beta$  the regularization parameters or hyper-parameters. The second term on the right hand side of equation,  $\pm E_w$ , is the *weight decay* and with  $\pm$ , the weight decay rate, favors small values of  $w$  and decreases the tendency of a model to overfit. Large values of  $\alpha$  lead to posterior densities of weights that are highly concentrated around zero, so that the weights effectively disappear discounting connections in the network. If  $\alpha \ll \beta$  then the training algorithm will make the errors smaller. If  $\alpha \gg \beta$ , training will emphasize weight size reduction at the expense of network errors, thus producing a smoother network response [12].

After the data is taken, the density function for the weights can be updated according to Bayes' rule. The posterior distribution of  $w$  given  $\alpha$ ,  $\beta$ ,  $D$ , and  $M$  is:

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M) \cdot P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \quad (6)$$

where  $D$  is the training data set and  $M$  is the specific functional form of the neural network architecture considered.  $P(w|D, \alpha, \beta, M)$  is the posterior probability of  $w$  and  $P(D|w, \beta, M)$  the likelihood function which is the probability of the occurrence, giving the network weights.  $P(w|\alpha, M)$  is the prior distribution of weights under  $M$ ,  $P(D|\alpha, \beta, M)$  is a normalization factor or evidence for hyperparameters  $\alpha$  and  $\beta$ .

We assume that the noise in the training set data is Gaussian and that the prior distribution for the weights is Gaussian, then probability densities write:

$$\begin{aligned} P(D|w, \beta, M) &= \left(\frac{\beta}{\pi}\right)^{n/2} \exp(-\beta E_D), \\ P(w|\alpha, M) &= \left(\frac{\alpha}{\pi}\right)^{N/2} \exp(-\alpha E_w), \end{aligned} \quad (7)$$

where  $n$  and  $N$  are the number of observations and total number of network parameters, respectively. We substitute these two probabilities in the equation 6, we obtain:

$$\begin{aligned} P(w|D, \alpha, \beta, M) &= \frac{\left(\frac{\beta}{\pi}\right)^{n/2} \cdot \left(\frac{\alpha}{\pi}\right)^{N/2} \exp(-(\beta E_D + \alpha E_w))}{P(D|\alpha, \beta, M)} \\ &= \frac{1}{Z_F(\alpha, \beta)} \cdot \exp(-F(w)) \end{aligned} \quad (8)$$

Maximizing the posterior probability  $P(w|D, \alpha, \beta, M)$  is equivalent to minimizing the regularized objective function  $F = \beta E_D(D|w, M) + \alpha E_w(w, M)$ .

Considering the joint posterior density by:

$$P(\alpha, \beta|D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta, M)}{P(D|M)} \quad (9)$$

Now the equation 6 is, according to Mckay 1992 [23] :

$$\begin{aligned} P(D|\alpha, \beta, M) &= \frac{P(D|w, \beta, M) \cdot P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \\ &= \frac{Z_F(\alpha, \beta) \cdot \exp(-\beta E_D - \alpha E_w)}{\left(\frac{\beta}{\pi}\right)^{n/2} \cdot \left(\frac{\alpha}{\pi}\right)^{N/2} \cdot \exp(-F(w))} \\ &= \frac{Z_F(\alpha, \beta)}{\left(\frac{\beta}{\pi}\right)^{n/2} \cdot \left(\frac{\alpha}{\pi}\right)^{N/2}} \end{aligned} \quad (10)$$

where  $Z_F(\alpha, \beta)$  can be estimated by Taylor series expansion: see (Foresee et al 1997 [12]). The objective function  $F(w)$  has the shape of a quadratic in the neighborhood of the minimum point, then  $F(w)$  is expanded around the minimum point of the posterior density, where the gradient is zero.

$$Z_F(\alpha, \beta) \approx (2\pi)^{N/2} (\det((\mathbf{H}^{MP})^{-1}))^{1/2} \exp(-F(w^{MP})) \quad (11)$$

where  $\mathbf{H} = \beta \nabla^2 E_D + \alpha \nabla^2 E_w$ , the Hessian matrix of the objective function. Values of regularization parameters,  $\alpha$  and  $\beta$  are calculated as:

$$\alpha^{MP} = \frac{\gamma}{2E_w(w^{MP})} \quad \text{and} \quad \beta^{MP} = \frac{n - \gamma}{2E_D(w^{MP})} \quad (12)$$

with  $\gamma = N - 2\alpha^{MP} \text{Tr}(\mathbf{H}^{MP})^{-1}$ , the effective number of parameters, and  $N$  the total number of parameters in the network. The Bayesian optimization of the regularization parameters requires the computation of the Hessian matrix of the function  $F(w)$  at the minimum point  $w^{MP}$  [12]. Mackay 1992 proposes an approach in [23]: the Gauss-Newton approximation to the Hessian matrix can be used if the Levenberg-Marquardt optimization algorithm is employed to locate the minimum point.

2) *Levenberg-Marquardt optimization*: The Levenberg-Marquardt algorithm is a robust numerical optimization technique for mapping as well as function approximation. We define the least squares cost function  $J(w)$  by

$$J(w) = \frac{1}{2} \sum_{i=1}^n (y'_i - y_i)^2 \quad (13)$$

its gradient is therefore defined by the vector

$$\nabla J(w) = \left( \frac{\partial J}{\partial w_1}, \frac{\partial J}{\partial w_2}, \dots, \frac{\partial J}{\partial w_n} \right)^T \quad (14)$$

The Hessian matrix of the cost function, and has the form:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 J(w)}{(\partial w_1)^2} & \frac{\partial^2 J(w)}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 J(w)}{\partial w_1 \partial w_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 J(w)}{\partial w_n \partial w_1} & \frac{\partial^2 J(w)}{\partial w_n \partial w_2} & \dots & \frac{\partial^2 J(w)}{(\partial w_n)^2} \end{pmatrix}$$

The Levenberg-Marquardt algorithm, which also belongs to the class of quasi-Newtonian methods, obeys the following formula for updating the parameters at  $l$  iteration:

$$w^{l+1} = w^l - [H(w^l) + \mu_{l+1}I]^{-1} \nabla J(w^l) \quad (15)$$

where  $\mu_{l+1}$  is Levenberg's damping factor, which is adjusted at each iteration and guides the optimization process, and  $I$  is the identity matrix. We will find in [34] a popular alternative to the Gauss-Newton method of finding the minimum of a function.

From a practical point of view, the Bayesian approach to neural networks brings important improvements: as all calculations are done from the training base, it is no longer necessary to have a validation base. It is therefore possible to use all the data available to estimate the weights of the network.

3) *Regularized cost function and BFGS algorithm*: Another method of regularization consists in modifying the performance function, we add a penalizing term consisting of the mean of the sum of squares of the network weights to the cost function.

$$F = \frac{1}{2} \left( \frac{1}{N} \sum_{i=1}^n (y'_i - y_i)^2 \right) + \frac{1}{2} \left( \frac{1}{n} \sum_{i=1}^n w_i^2 \right). \quad (16)$$

We implement with this regularized function the BFGS algorithm.

This algorithm (named after its inventors Broyden, Fletcher, Glodfarb and Shanno) is based on an approximation of Newton's method. The parameter update rule is defined as follows:

$$w^{l+1} = w^l - \mu_{l+1} M_{l+1} \nabla F(w^l) \quad (17)$$

where  $M_{l+1}$  is an iteratively calculated approximation of the inverse of the Hessian matrix. The approximation of the inverse of Hessian is modified at each iteration according to the following rule:

$$M_{l+1} = M_l + \left( 1 + \frac{\gamma_l^T M_l \gamma_l}{\delta_l^T \gamma_l} \right) \frac{\delta_l^T \delta_l}{\delta_l^T \gamma_l} - \frac{\delta_l \gamma_l^T M_l + M_l \gamma_l \delta_l^T}{\delta_l^T \gamma_l} \quad (18)$$

where  $\gamma_l = \nabla F(w_l) - \nabla F(w_{l-1})$  and  $\delta_l = w_l - w_{l-1}$ . The initial value of the matrix  $M$  is generally the identity matrix, value to which  $M_{l+1}$  will also be reset during the algorithm if it turns out to be no longer definite positive.

The interest of the BFGS algorithm lies in that it makes it possible to be freed from the computation of the inverse of the Hessian matrix (which can itself prove to be delicate in certain cases), by iteratively estimating an approximation of this inverse matrix according to formula (13). This quasi-Newtonian method is only effective near the minimum of the cost function.

#### IV. RESULTS AND DISCUSSIONS

The objective is therefore to find a compromise between the quality of learning and the capacity for generalization. Also first, we evaluate training performance of the Levenberg-Marquardt(LM) algorithm without regularization, compared with two regularization algorithms. This part presents the results of the different stages of the modeling: a phase of calculation carried out in order to determine the architecture of the optimal ANN and its validation before moving on to predictions. Secondly, regularized ANN were employed to predict strain of the polymer B and thirdly the polymers A-150, A-185.

A. Implementation of ANN

We trained the LM, BR and BFGS networks, multiple times with 33 vectors by changing the number of hidden layer neurons and selected the one which gave best results for prediction. For the Bayesian Regularization Artificial Neural Network (BRANN), the regularization parameters  $(\beta, \alpha)$  were optimized following Equ 5. The validation set is not essentially required in the case of the regularization methods.

To evaluate the performance of the three networks, mean squared error (MSE) and correlation coefficient R are estimated. The MSE measures the deviation between original values and predicted ones, and R provides information on the strength of correlation between them. They are defined by the equations below:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2 \quad (19)$$

$$R = \frac{\sum_{i=1}^N (y_i - \bar{y})(y'_i - \bar{y}')}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (y'_i - \bar{y}')^2}} \quad (20)$$

where  $y_i$  is the observed value and  $y'_i$  is the network output value.  $y'$  and  $\bar{y}'$  are respectively the average of the real value and the network output value, and  $N$  is the sample number. The MSE and the number of neurons in the hidden layer are investigated to construct the optimal structure of the neural network.

The MSE depending on iteration (epochs) of the BRANN is shown in the Figures above. In red, the error on the test set, in green the error on the validation set. In blue, the learning error. If the validation and test error increase while the training error continues to decrease then there is overfitting. The training stopped when the MSE value was achieve. A negligible value of the MSE indicates the high degree of correlation among input variables. Training stops when any of these conditions occurs: 1) the maximum number of epochs is reached, 2) performance of the network with the number of neurons has met a suitable level, 3) performance is minimized to the goal 4) the gradient was below a suitable target 5)  $\mu$  exceeds  $\mu_{max} = 10^{10}$ . In conclusion, we see the effectiveness of the regularization methods. Indeed, we clearly observe in the left figure 13 of the supervised training performance an overfitting.

The two figures of histograms represent the errors between target values and predicted values after training ANN. These error values indicates how predicted values are differing from the target values. Y-axis represents the number of samples from dataset, which lies in a particular bin. For example for the BRANN, we have a bin corresponding to the error of -0.0031 and the height of that bin for training dataset lies near to 11, the height of bar in the bar plot means how many data points are near the bin value. It means that 11 samples from training dataset have an error lies in the following range. Zero error line corresponds to the zero error value on the X-axis. We can see on the left histogram that the errors are much more dispersed: the difference is between -1.95 and 1.05, the majority of errors correspond to 7 samples with an error of -0.057.

The figures of the regression coefficient show us a better performance of R for the BRANN with regularization whose

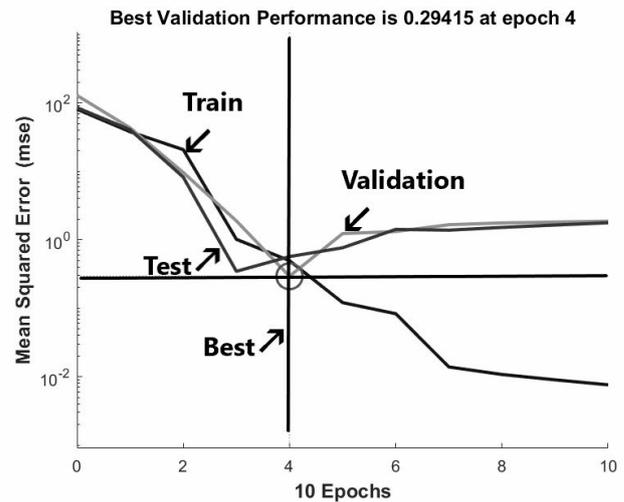


Figure 12. Performance of the feedforward neural network without regularization

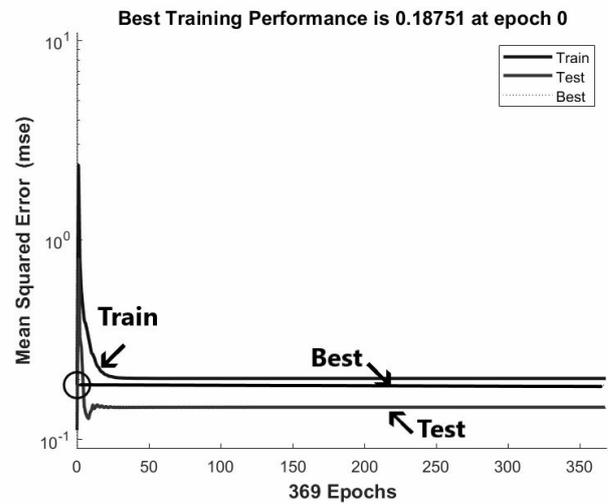


Figure 13. Performance of the Bayesian neural network

R is close to 0.95 against 0.92 for the NN without regularization. For the BRANN, the outputs are correlated with the corresponding target values for training and testing, the R value is 0.9599 for the total response. There is a relatively linear relationship between outputs and targets. These results show a good fit at the level of training and testing. It was shown that there is a good correlation between the predicted values of BRANN and the experimental values. After several tests the optimal neural architecture is composed of 12 neurons with the performance equal to 0.1857.

Now the optima structure and the performance parameters of the three networks are summarized Table III.

Table III  
MEAN SQUARED ERROR AND EPOCHS FOR DIFFERENT ARCHITECTURES.

Network models	Network structures	MSE	Epochs
LM	4 - 5 - 1	0.4642	8
BFGS	4 - 12 - 1	0.1861	300
BR	4 - 12 - 1	0.1803	369

The models training is performed with 4 to 15 hidden neurons. By increasing the neural network structure of one neuron each time and comparing the mean square error, it

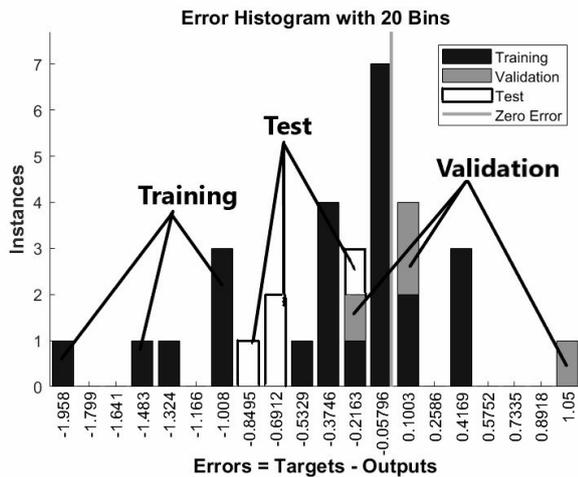


Figure 14. Error histogram for the neural network without regularization

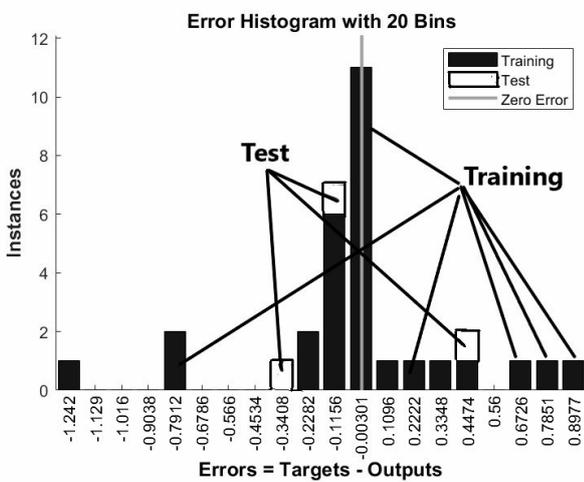


Figure 15. Error histogram for the Bayesian regularization neural network

attains a minimum for the BRANN model with 12 neurons. The minimum values are obtained for the two regularized models. So, the model can be explained well with small data set when Bayesian Regularization is used for training the networks. Once the BRANN model has learned well, it interpolates the data according to an implicit function of the following form:

$$\varepsilon = f(t, T, E, \sigma) \tag{21}$$

**B. ANN prediction results for the polymer B.**

Now, we are going to make all of our predictions in our work with the BR and BFGS algorithms.

In this section, first we trained the models with a dataset containing 32 samples (at 33 samples we remove the last value for aging time t=15th month for  $T = 70^{\circ}C$ ) of the polymer B. After having selected for each model the optimal structure we predicted the following value for  $T = 70^{\circ}C$  and aging time t = 15th month.

In order to verify the ability of the models to predict data outside the database, some following data for aging time t=15 months. The prediction results and the statistical parameters are presented in Table IV. It can be seen that the regularized models have good performance. This means

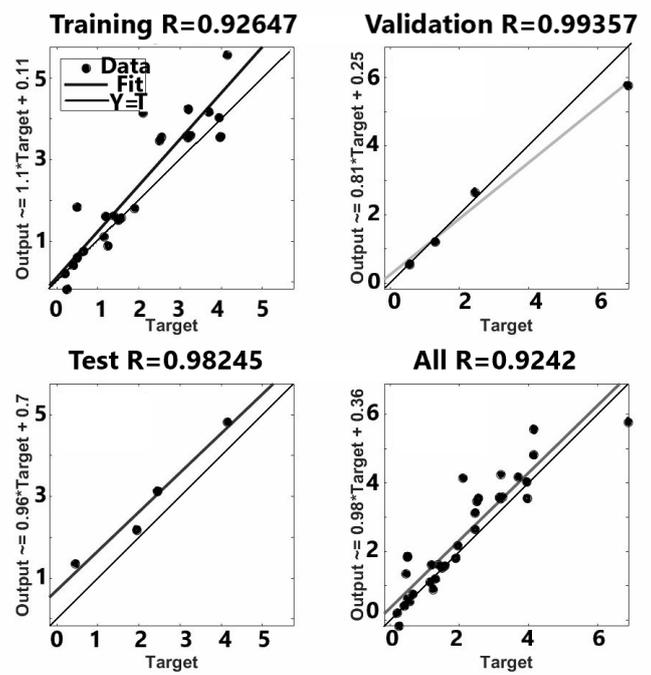


Figure 16. Regression of the neural network without regularization

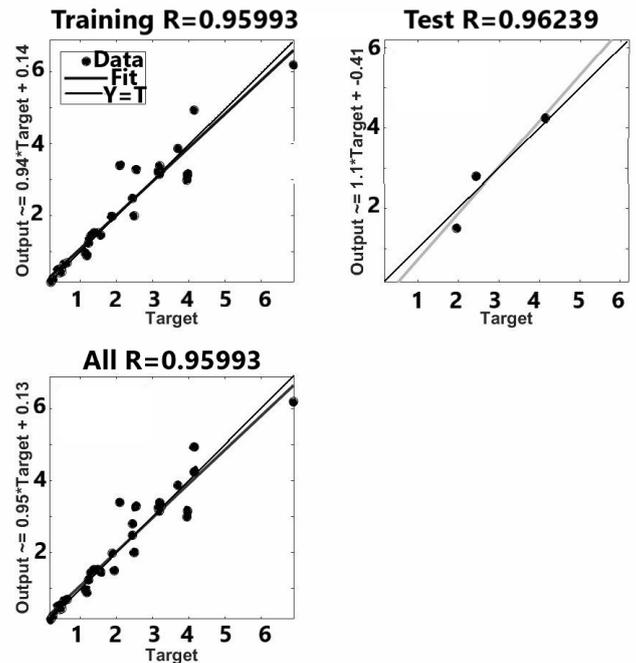


Figure 17. Regression of Bayesian regularization neural network

that the regularized models have a good ability to predict the unknown data and a better performance compared to the unregulated model.

According to the performance calculations for the strain prediction for  $T = 70^{\circ}C$ , and aging time t=15 months, Bayesian regularization gives the best result with a mean square error of 0.0036 against 0.0043 for BFGS.

Secondly, a case with even smaller subset of the data was considered with 28 samples (at 33 samples we remove the last five samples for aging time 3th, 6th, 9th, 12th and

Table IV  
STATISTICAL PARAMETERS OF THE PREDICTED STRAIN OF THE POLYMER B FOR  $T = 70^{\circ}C$ , T=15 MONTHS.

Network models	$\epsilon(\%)$	predicted $\epsilon(\%)$	MSE
BFGS	1.45	1.38	0.0043
BR	1.45	1.51	0.0036

15th months for  $T = 70^{\circ}C$ ) and we predicted strain values for aging time the 3th, 6th, 9th, 12th and 15th months and  $T = 70^{\circ}C$ .

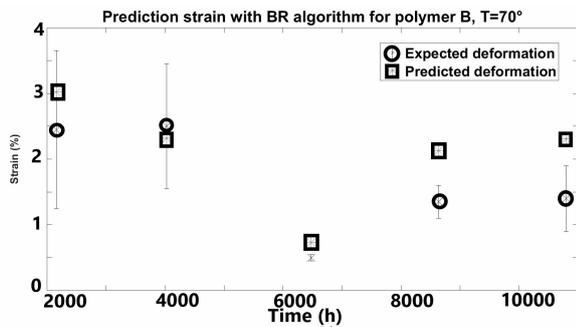


Figure 18. Prediction by the BR algorithm of the deformation of the polymer B, for the last five terms of the experiment

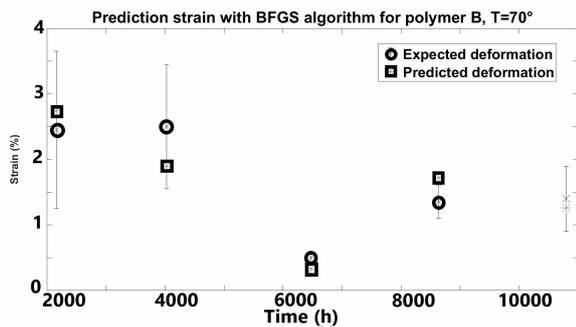


Figure 19. Prediction by the BFGS algorithm of the deformation of the polymer B, for the last five terms of the experiment

As shown in Fig. 19, with the BFGS algorithm, the last five values predicted of strain are close to the desired outputs results, in particular within the interval of uncertainties. These results show that the BFGS regularization is accurate even with a small dataset. However, for BRANN the predicted strain values show deviation from the desired target result and three points do not belong to the uncertainty interval. The MSE corresponding are given Table V and it confirms that the performance of BRANN is a little worse than that of BFGS:

Table V  
PREDICTION OF THE STRAIN OF THE POLYMER B FOR  $T = 70^{\circ}C$ , AND AGING TIME T= 3, 6, 9, 12 AND 15 MONTHS.

Network models	MSE
BFGS	0.108
BR	0.356

From the cases, we observed that the predictions of BRANN model are closer to the experimental target values even with a small training data set. But in the last case where

a prediction of the strain is requested for many aging times  $t=3, 6, 9, 12$  and 15 months, BFGS performs better than BR. The accuracy of the regularized models decrease when increasing the number of points to predict. Both regularized algorithms are able to predict, despite a small data set, for aging times depending on the mechanical characteristics of a polymer. This means that during experimental work a neural network is able to predict for the next three, six, nine months.

This type of prediction, with a quality of precision even for small datasets which is often the case in the field of materials, can make it possible to continue and complete experiments. It allows to reduce the time and the cost in long experiments, for the study of the durability of polymers. Monitoring the aging of polymers over several months is costly and restrictive. ANNs can be an effective tool, even on small samples, to test the behavior and performance of materials.

C. Prediction of the strain for the polymers A-150 and A-185.

Now, an already trained neural network is used to make predictions on new data with variables that have been generated by the same underlying processes and relationships as the original dataset that was used to train the model. The possibility of generalization is an essential characteristic of neural networks. We tested Bayesian methods and regularized BFGS methods on other polymer systems at the same conditions that the system B. We explore characterization of polymer systems aged in traction on systems A-185 and A-150. The entire sample of polymer B was used to train the neural network. Time aging, temperature, elasticity modulus and stress was used as inputs contained 33 experimental data and the strain as output. Then, the trained neural network on polymer B was requested to predict the points of the strain for  $T = 70^{\circ}C$  of polymers A - 150, A- 185. A comparison is then made between the forecast values obtained with the experimental data. Figures 20, 21, 22 and 23 below next illustrate this comparison. The red point represents the expected results according to the network formed and the blue point represents the experimental results.

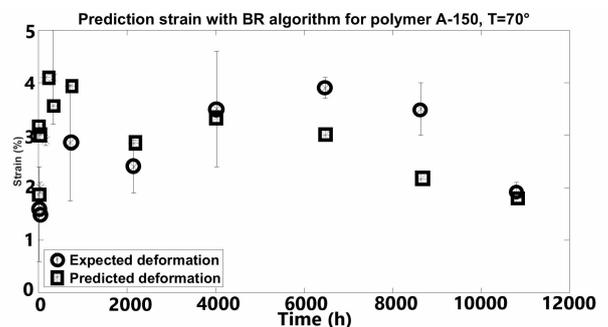


Figure 20. The comparison of experimental and predicted strain for polymer A-150 with Bayesian regularization

The table below gives MSE corresponding of the prediction of the ANNs on these polymer systems.

A perfect match between the experimental values and the predicted values of the polymers A was not expected. According to the results of Table VI, the performance of the

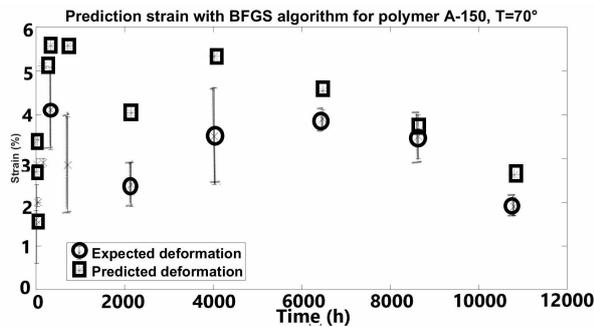


Figure 21. The comparison of experimental and predicted strain for polymer A-150 with BFGS algorithm

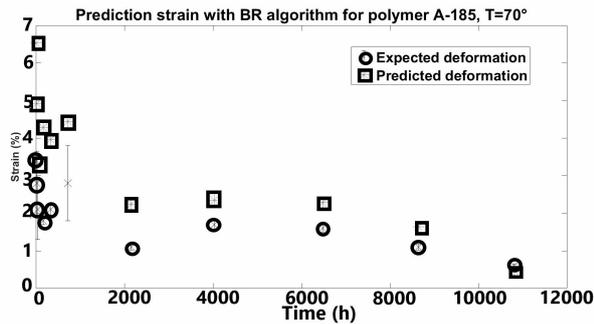


Figure 22. The comparison of experimental and predicted strain for polymer A-185 with Bayesian regularization

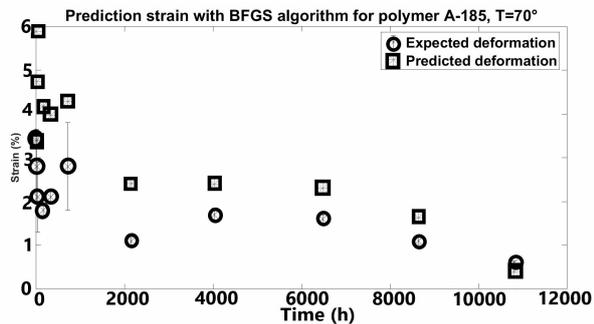


Figure 23. The comparison of experimental and predicted strain for polymer A-185 with BFGS algorithm

Table VI  
PREDICTION OF THE STRAIN FOR  $T = 70^{\circ}C$  OF THE POLYMERS A-150 AND A-185.

Network models	MSE (A-185 $70^{\circ}C$ )	MSE (A-150 $70^{\circ}C$ )
BFGS	2.35	0.86
BR	2.4	2.45

network are a little worse than the prediction on polymer B, with however a larger number of points to be predicted. We note, on the one hand, that the BFGS algorithm gives better results than BR for the two polymers A. On the other hand, predictions are better for the long times and much more accurate than on short times. Here we therefore reach the limits of the neural network training on a small set of data. However, we note that the predicted strains have the same tendency as the experimental strains.

D. Experimental data of the strain for the polymers A-150 and A-185 and comparison with predicted data.

Regarding the properties of polymers A-185, from the experimental data, we see Figure 24 evolution of the deformation which seems to take place in three stages. First, a decrease in properties during the first week, then they increase for up to 1 month to decrease until the end of aging

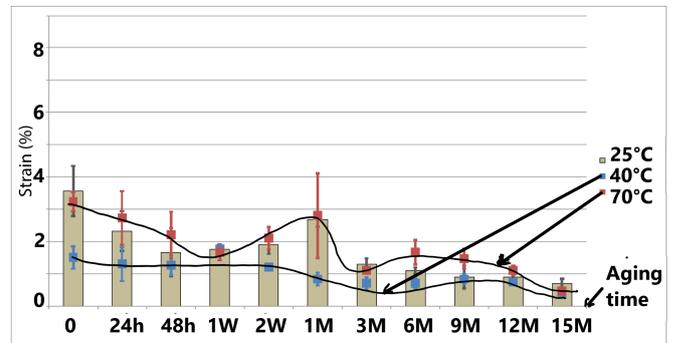


Figure 24. Evolution of the deformation of the A-185 system as a function of the aging time and the test temperature.

The initial oxidation of the samples seems to lower the properties at first, then the strong water absorption plasticizes the network and improves the mechanical properties up to 1 month of aging. Subsequently, oxidation is predominant, totally weakening the test pieces and the polymer network loses completely its mechanical properties. Tests carried out at  $-40^{\circ}C$  weaken the structure even more due to the drop in properties, while tests at  $70^{\circ}C$  show the same evolution as at room temperature with a slightly more ductile network.

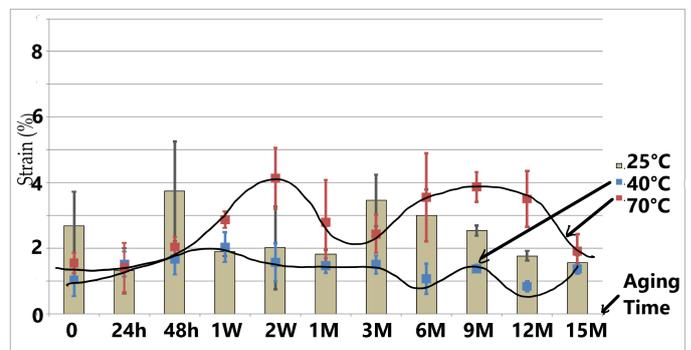


Figure 25. Evolution of the deformation of the A-150 system as a function of the aging time and the test temperature.

Figure 25 shows the evolution of the deformation of the A-150 system as a function of the duration of aging and the test temperature, respectively. The values of the deformation constantly oscillate between 0 and one month. With regard to elongation at break, no trend was noted up to 3 months of aging. Then, the values decrease during aging. As the polymer A-150 contains a soluble phase, during the first stages of aging, there is a competition between the diffusion of water and the loss of low masses by leaching observed during the rheological study. So the superposition of these two phenomena generate opposite consequences: a water intake leads to a reduction in stresses and an increase in deformations while a loss of small masses, which play the

role of plasticizer, leads to an increase in stresses and a reduction in deformations because we have rigidification of the structure. Beyond one month of aging, the oxidation highlighted above weakens the polymer network, which sees its properties drop, but the values reached are still higher than those obtained for polymer A-185. We find a brittle system at  $-40^{\circ}\text{C}$  and ductile at  $70^{\circ}\text{C}$  with values of stress and strain at break always higher than those of A-185.

The study of the variation of the mechanical properties in traction and by determination of the resilience, shows during the first month of aging, a plasticization of systems B and A-185. There follows a collapse of the mechanical properties after 3 months of aging due to the weakening of the systems generated by oxidation and hydrolysis during aging. It is all the same interesting to note that despite a non-optimized crosslinking, the mechanical characteristics of the polymer A-150, which were initially lower, are superior to those of the two other systems after 1 year of aging.

## V. CONCLUSION

In with work, Bayesian Regularization and BFGS with modified performance function models are employed to predict the strain of traction-aged polymer systems. In the first stage, regularized ANN and ANN model without regularization were built using a very small dataset of 33 samples. The training phase of the ANNs is performed taking into consideration several parameters, such as the aging time, temperature, tensile stress and Young's modulus. The optimal architecture model which contains a sigmoid function and an output layer which contains a linear function, is evaluated using mean square error (MSE) and the regression value R. It is concluded that the Bayesian regularization training algorithm and BFGS regularized algorithm show better performance than the Levenberg-Marquardt algorithm without regularization. Regularized algorithm can solve the overfitting problem which is not the case of Levenberg-Marquardt algorithm for a small data set.

In a second stage, regularized ANN models, with three datasets of 32, 30 and 28 samples, are used to predict the percent strain of the system B for several aging times. BRANN showed higher performance for one aging time, three aging times predictions. Finally in a third stage, these regularization methods are also used for two different polymer systems : the polymers A-150 and A-185 that were not in the training dataset. Considering the limitations of the model, due to the small dataset, the neural network can more accurately predict the results of the two polymers for long times than for short times. However, the network successfully predicted the strain trend for both polymers.

This allows us to conclude that these two regularized ANN are reliable despite small data. These ANN approaches can be used to predict the trend for the next few months, which saves time and cost in the experimental field.

## REFERENCES

- [1] Helene Canot, Philippe Durand, Emmanuel Frenod, Bouchera Hassoune-Rhabbour and Valerie.Nassiet, "Regularized Artificial Neural Networks for Predicting the Strain of Traction-Aged Polymer Systems Part II", Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2022, 6-8 July, 2022, London, U.K., pp. 15-20
- [2] OT. Adesina, T. Jamiru, I. Daniyan, R. Sadiku, O. Ogunbiyi, O. Adesina, LW.Beneke (2020). Mechanical property prediction of SPS processed GNP/PLA polymer nanocomposite using artificial neural network. Cogent Engineering. 7. DOI:10.1080/23311916.2020.1720894.
- [3] E. Arzaghi, MM. Abaei, R. Abbassi, V. Garaniya, C. Chin, F. Khan (2017) Riskbased maintenance planning of subsea pipelines through fatigue crack growth monitoring. Eng Fail Anal 79, pp. 928-939.
- [4] I. BabuÅjka, Z. Sawlan, M. Scavino, B. Szabo, R. Tempone. (2016) Bayesian inference and model comparison for metallic fatigue data. Comput Methods Appl Mech Eng 304, pp. 171-196
- [5] L. Barbosa, G. Gomes, A. Ancelotti. (2019). Prediction of temperature-frequency-dependent mechanical properties of composites based on thermoplastic liquid resin reinforced with carbon fibers using artificial neural networks. The International Journal of Advanced Manufacturing Technology. 105, pp. 1-14. 10.1007/s00170-019-04486-4.
- [6] H E. Balcioglu, A. Agdas Seckin and M. Aktas. (2016) Failure load prediction of adhesively bonded pultruded composites using artificial neural network. Journal of Composite Materials 2016, Vol. 50(23), pp. 3267-3281. DOI: 10.1177/0021998315617998
- [7] E. Burgaz, M. Yazici, M. Kapusuz, S. H. Alizir, H. Åzcan. (2014). Prediction of thermal stability, crystallinity and thermomechanical properties of poly (ethylene oxide)/clay nanocomposites with artificial neural networks. Thermochemica Acta. 575,pp. 159-166. 10.1016/j.tca.2013.10.032.
- [8] D. Colombini, J.J. Martinez-Vega, and G. Merle. (2002) Dynamic mechanical investigations of the effects of water sorption and physical ageing on an epoxy resin system. Polymer, 43(16), pp. 4479-4485.
- [9] A.Doblies, B. Boll, and B. Fiedler. (2019) Prediction of Thermal Exposure and Mechanical Behavior of Epoxy Resin Using Artificial Neural Networks and Fourier Transform Infrared Spectroscopy.Polymers (Basel). Feb; 11(2), 363.
- [10] T. Dyakonov, P.J. Mann, Y. Chen, and W.T.K. Stevenson. (1996) Thermal analysis of some aromatic amine cured model epoxy resin systems. II: Residues of degradation. Polymer Degradation and Stability, 54(1), pp. 67-83.
- [11] B.C. Ennis, P.J. Pearce, and C.E.M. Morris. (1989) Aging and performance of structural film adhesives. III. Effect of humidity on a modern aerospace adhesive. Journal of Applied Polymer Science, 37(1), pp. 15-32.
- [12] F. Dan. Foresee and Martin T. Hagan. (1997) "Gauss-Newton approximation to Bayesian learning." Proceedings of the International Joint Conference on NeuralNetworks, June.
- [13] L. Gavard, Hkdh Bhadeshia, D. J. C. MacKay, and S. Suzuki. (1996) Bayesian neural network model for austenite formation in steels. Materials Science and Technology, 12(6), pp. 453-463.
- [14] ATC. Goh. (1995) Back-propagation neural networks for modeling complex systems. Artificial Intelligence in

- Engineering 9, pp. 143-15.
- [15] N. Huber, Ch. Tsakmakis, (2001). A neural network tool for identifying the material parameters of a finite deformation viscoplasticity model with static recovery. *Computer Methods in Applied Mechanics and Engineering* 191, pp. 353-384.
- [16] Z. Zhang, P. Klein, and K. Friedrich. (2002) Dynamic mechanical properties of ptfе based short carbon fibre reinforced composites: experiment and artificial neural network prediction. *Composites Science and Technology*, 62(7-8), pp.1001-1009.
- [17] S. Koroglu, P. Sergeant, and N. Umurkan. Comparison of analytical, Finite element and neural network methods to study magnetic shielding. *Simul. Model. Pract. Theory*, 18(2):206216, 2010. doi:10.1016/j.simpat.2009.10.007.
- [18] N. Kiuna, C.J. Lawrence, Q.P.V. Fontana, P.D. Lee, T. Selerland, P.D.M. Spelt. A model for resin viscosity during cure in the resin transfer moulding. process. *Composites : Part A* (33), pp. 1497-1503, 2002.
- [19] M. Lefik and B.A. Schreer. (2003) Artificial neural network as an incremental non-linear constitutive model for a finite element code. *Comput. Methods Appl. Mech. Eng.*, 192(28-30):32653283, 2003. doi:10.1016/S0045-7825(03)00350-5.
- [20] S.Mahmoudi. (2017) Dynamique des structures composites linéaire et non-linéaire en présence d'endommagement. (Doctoral thesis). Université de Bourgogne Franche-Comté.
- [21] <https://www.mathworks.com/help/nnet/ref/trainbr.html>
- [22] W. S. McCulloch W. Pitts. (1943) A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115-133.
- [23] D.J.C. Mackay. (1992) Bayesian interpolation. *Neural Comput.* 4, pp. 415-447.
- [24] I. Merdas, F. ThomINETTE, A. Tcharkhtchi, and J. Verdu. (2002) Factors governing water absorption by composite matrices. *Composites Science and Technology*, 62(4), pp. 487-492.
- [25] D. W. Marquardt,(1963), An algorithm for least-squares estimation of non-linear parameters, *Journal of the society for Industrial and Applied Mathematics*, 11(2), pp. 431-441.
- [26] M. Minsky S. P. Perceptrons. (1969) an Introduction to Computational Geometry Cambridge Ma.
- [27] L. Qingbin, J. Zhong, L. Mabao, W. Schichum, (1996). Acquiring the constitutive relationship for a thermal viscoplastic material using an artificial neural network. *Journal of Materials Processing Technology* 62, pp. 206-210.
- [28] Y. Pei, K. Wang, M. Zhan, W. Xub, X. Ding. Thermal-oxidative aging of DGEBA/EPN/LMPA epoxy system : Chemical structure and thermomechanical properties. *Polymer degradation and stability*, 96:1179-1186, 2011.
- [29] S. Popineau, C. Rondeau-Mouro, C. Sulpice-Gaillet, and M.E.R. Shanahan. (2005) Free/bound water absorption in an epoxy adhesive. *Polymer*, 46(24), pp. 10733-10740.
- [30] L.Poussines. (2012) Développement de nouveaux matériaux pour l'infusion de composites. PhD . Spécialité : science et génie des matériaux.
- [31] S. Pruksawana, G. Lambard, S. Samitsu, K. Sodeyama and M. Naitoa. Prediction and optimization of epoxy adhesive strength from a small dataset through active learning. *SCIENCE AND TECHNOLOGY OF ADVANCED MATERIALS* 2019, VOL. 20, NO. 1, pp. 1010-1021 <https://doi.org/10.1080/14686996.2019.1673670>
- [32] F. Rosenblatt. (1962) Principles of neurodynamics.
- [33] S. B. Singh, Hkdh Bhadeshia, D. J. C. MacKay, H. Carey, and I. Martin. (1998) Neural network analysis of steel plate processing. *Ironmaking Steelmaking*, 25(5), pp. 355-365.
- [34] D.C. Souza, (2015) Neural Network Learning by the Levenberg-Marquardt Algorithm with Bayesian Regularization. Available online: <http://crsouza.blogspot.com/feeds/posts/default/webcite>.
- [35] G.Z. Xiao and M.E.R. Shanahan. (1998) Swelling of DGEBA/DDA epoxy resin during hygrothermal ageing. *Polymer*, 39(14), pp. 3253-3260.
- [36] J. Verdu. (2000) Action de l'eau sur les plastiques. *Techniques de l'ingénieur*, pp. 1-8.
- [37] M. Wiciak-PikuÅa, A. Felusiak-Czyryca and P. Twardowski (2020) Tool Wear Prediction Based on Artificial Neural Network during Aluminum Matrix Composite Milling. MDPI. Proceedings of the 2020 IEEE International Workshop on Metrology for AeroSpace, Pisa, Italy, 22-24 June 2020.
- [38] Zhang, Zhong Friedrich, Klaus. (2003). Artificial Neural Networks Applied to Polymer Composites: A Review. *Composites Science and Technology - COMPOSITES SCI TECHNOL.* 63. 2029-2044. 10.1016/S0266-3538(03)00106-4.