Regularized Artificial Neural Networks for Predicting the Strain of Traction-aged Polymer Systems Part I

Helene Canot, Philippe Durand, Emmanuel Frenod, Bouchera.Hassoune Rhabbour, Valerie Nassiet

Proceedings of the World Congress on Engineering 2022
WCE 2022, July 6 - 8, 2022, London, U.K.

Regularized Artificial Neural Networks for Predicting the Strain of Traction-aged Polymer Systems Part I

Helene Canot, Philippe Durand, Emmanuel Frenod, Bouchera.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 = $-40°C$, $25°C$, $70°C$. 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. This paper is considered a first part of the problem. In this part, we give the setting of the experimental problem and we approach the theoretical part concerning Bayesian regularization and the BFGS algorithm. The article of part II will present the numerical results as well as its analysis.**

***Index Terms*—rtificial Neural Network Multi-Layer Perceptron Bayesian regularization Levenberg Marquardt and BFGS Polymer system Lifetimertificial Neural Network Multi-Layer Perceptron Bayesian regularization Levenberg Marquardt and BFGS Polymer system LifetimeA**

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. In this paper, 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 $T = 150°C$ for two hours with a conversion rate of 89% (noted A-150), system A polymerized at $T = 185°C$ 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 $T = 70°C$. 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 Génie 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 [6]), by degradation of the polymer by hydrolysis (Ennis et al. 1989 [9]); (Xiao et al. 1998 [33]), by differential swelling linked to concentration gradients (Merdas et al. 2002) [22] 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 [8]); (Popineau et al. 2005 [27]). 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 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 systems through the use of multi-layer perceptron (MLP) type with gradient backpropagation. In recent years, many researchers have used ANN in the field of composite materials to predict their behavior Zhang (2003) [36] and Goh (1995) [12]. Qingbin et al. (1996) [25] 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

Helene Canot is in Université de Bretagne-Sud, UMR 6205, LMBA, F-56000 Vannes, France e-mail:helene.canot@univ-ubs.fr,
Philippe Durand is in M2N, Conservatoire National des Arts et Métiers, 292 rue Saint Martin, 75141 Paris FRANCE e-mail:philippe.durand@lecnam.net,
Emmanuel Freod is in Université de Bretagne-Sud, UMR 6205, LMBA, F-56000 Vannes, France e-mail:emmanuel.frenod@univ-ubs.fr ,
Valerie Nassiet and Bouchera.Hassoune-Rhabbour are in: ENIT, 47, avenue d'Azereix - BP 1629 - 65016 Tarbes CEDEX, France, Equipe IMF, France, e-mail:valerie.nassiet@enit.fr, bouchra.hassoune-rhabbour@enit.fr

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

network. It was able to approximate the constitutive relation for a thermal viscoplastic material. Huber and Tsakmakis (2001) [13] 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) [18] 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) [5] 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) [7] 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 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) [3] to model the temperature-frequency dependence of dynamic mechanical of thermoplastic polymers of advanced composites. They studied a new Elium® acrylic matrix developed by Arkema to evaluate the accelerated test methodology based on time-temperature superposition principle of Carbon Fiber/Elium® 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) [1] examine the potential of ANN for the prediction of mechanical properties, namely density and hardness of graphene nanoplatelet (GNP)/polylactic acid (PLA) nanocomposite developed under 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 developments have gone through three periods of activity. The first period in the 1940s was due to the work of McCulloch and Pitts (1943) [20]. The second occurred in the 1960s with Rosenblatt's perception theorem of perceptual convergence (1962) [30] and the work of Minsky and Papert (1969) [24] 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 ANN 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 has been employed to study various problems such as constitutive modeling : in [17] (2003) M. Lefik and B.A. Schreer. proposed a Bayesian Regularization Artificial Neural Network (BRANN) as a tool for numerical modelling of the constitutive behaviour of a physically non-linear body, or to study magnetic shielding in [15] 2010, where analytical, finite element and BRANN methods was compared to calculate the shielding efficiency of a cylindrical ferromagnetic shield. In [31] Singh et al. (1998) used a BRANN to predict the yield and tensile strength of rolled steel sheets as a function of chemical composition and processing parameters. Zhang et al. [14] (2002) employed BRANN to predict storage and loss modului of short fiber reinforced composites as a function of material composition and temperature. Gavard et al. [11] (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.

Arzaghi et al. [2] (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. [29] (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, H E. Balcıoglu et al. (2017) [4] used six different training artificial neural network algorithms such as Bayesian regularization, Levenberg–Marquardt and BFGS to predict the failure loads of bonded pultruded composite.

Recently, M. Wiciak-Pikuła et al. deals in [35] 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 imputs.

Our study has three steps. First, in part I, mechanical and physico-chemical characteristic of the polymer B and systems A-150, A-185 were extracted from the thesis work

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

of Laurence Poussines [28]. They are summarized below. Second, tensile tests were performed at three imposed temperatures by the industrialist, $-40°C$, $25°C$ and $70°C$. And thirdly in part II, 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, Youg 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 and the presentation of the Bayesian regularization, Levenberg–Marquardt and BFGS algorithms.

II. Material and methods

A. Material

In Laurence Poussines work [28], 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 $60°C$ and $+90°C$ 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] [16], The system A is a monocomponent. It consists of a stoichiometrics blends of a tetrafunctional epoxy prepolymer, the tetraglycidyl methylene dianiline (trademark TGMDA) and two hardeners, The 2,6-diethylaniline and le 2-isopropyl-6-methylaniline.

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 1

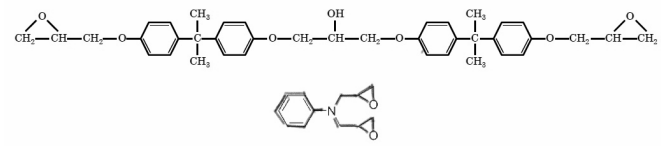

Fig. 1. Molecules present in resin B.

Two elements make up the hardener: 4,4-methanediyldicyclohexanamine and 3- (aminomethyl) -3, 5, 5-trimethylcyclohexanamine, the structures are shown schematically in Figure 2:

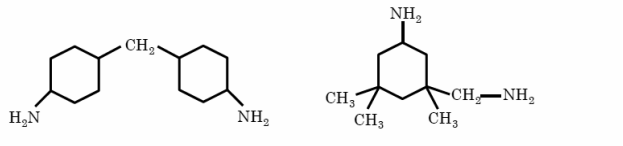

Fig. 2. Chemical structure of the substances present in the hardener.

B. Cured polymers'characterization before aging

The infusion of system A begins with preheating the resin to $80°C$. This then passes through a heating system maintained at $110°C$ to lower viscosity of system A up to 80 mPa.s then to diffuse in the preformed room maintained at $130°C$ in an oven or simply by heating lamps. Finally, the polymerization takes place at $150°C$ with these same heating systems for 2 hours. Adapting system A [28] 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.

An optimized polymerization cycle (2h at $185°C$ after a rate at $3°C$ /min) made it possible to obtain a rate of high conversion of 98% measured by IRTF as well as a glassy temperature Tg, determined by Differential Scanning Calorimetry (DSC), of $2113°C$. 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 $100°C$ + 3h at $140°C$ at $1°C$/min between isotherm) made it possible to obtain a high conversion rate of 98% measured by IRTF as well as a Tg of $1353°C$ measured by DSC. These properties are obtained without oxidation on the surface of the samples.

Tensile tests are carried out on $2x12x45mm^3$ rectangular-shaped samples. For each testing temperature imposed by the manufacturer ($40°C$, $25°C$, $70°C$ 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 0.2mm/min. System B behaves brittle at low temperatures and ductile at high temperatures whereas system A$185°C$ 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 $40°C$ and $70°C$.

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 $-40°C$, $25°C$ and $70°C$. The experimental data obtained from tensile tests for the B system studied at different temperatures and different aging times are shown in Figures 3, 4 and 5.

Tests carried out at room temperature show a slight decrease in Young's modulus up to one month of aging,

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

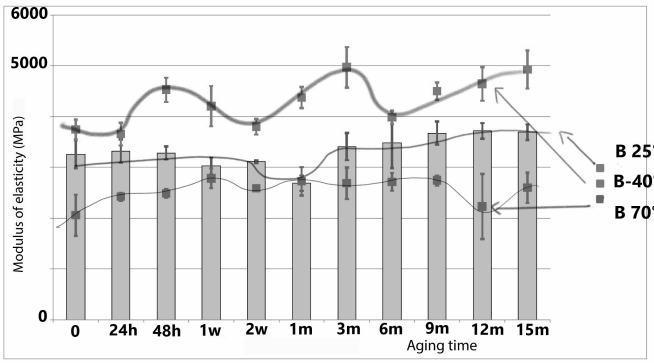

Fig. 3. Evolution of the elastic modulus of system B as a function of the aging time and the test temperature.

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.

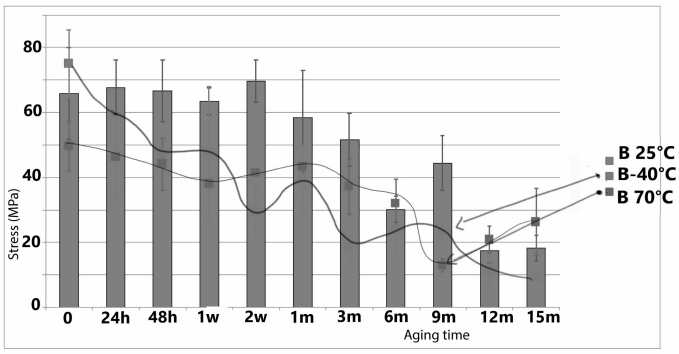

Fig. 4. Evolution of the breaking stress of system B as a function of the aging time and the test temperature.

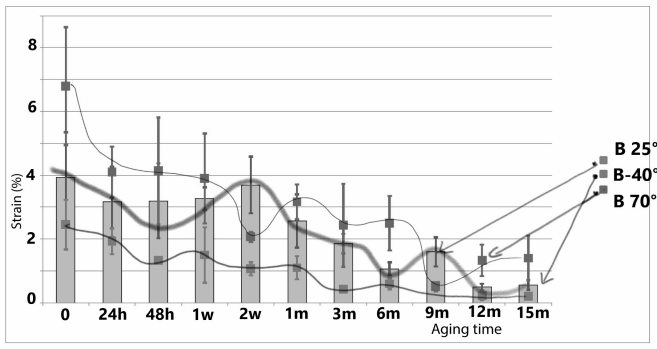

Fig. 5. Evolution of the deformation of system B as a function of the aging time and the test temperature.

The tests carried out at $-40°C$ 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 $70°C$, the modulus of elasticity and the tensile stress are lower and the deformation is higher than the values taken at $25°C$. The test pieces seem more ductile but a certain brittleness is notable due to the lower values of the 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 FTIR and hydrolysis with chains scissions leading to a mass loss during desorption test in oven at $70°C$.

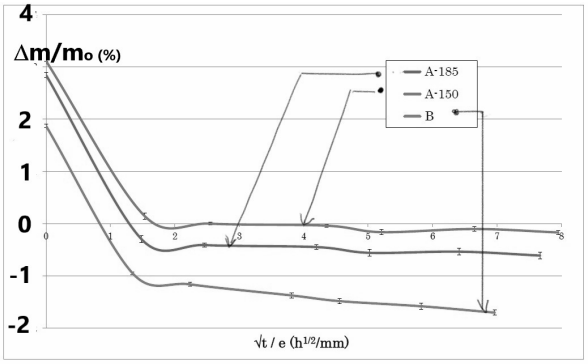

Fig. 6. Desorption curve at $140°C$ for the three systems.

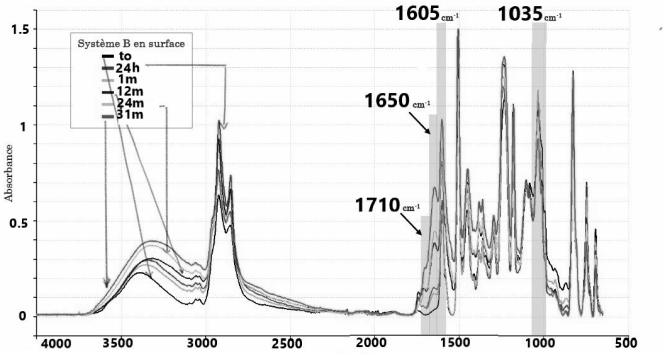

Fig. 7. Absorbance spectra of system B at the surface as a function of the number of waves (cm^{-1}) and as a function of the aging time.

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

D. Presentation of the database

In this study, an ANN was implemented with two different algorithms for regularization conditions and compared first

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

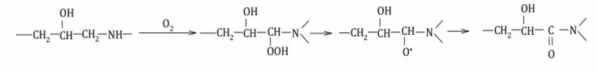

Fig. 8. Reaction mechanism of amide formation (Pei 2011) [26].

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 builded 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 I
VARIATION RANGE OF THE POLYMER B.

Time (months)	T $°C$	E (MPa)	$\sigma(MPa)$	$\varepsilon(\%)$
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 $70°C$ 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 for Levenberg-Marquard algorithm, trainbr for Bayesian Regularization algorithm, tansig, mapminmax and trainbfg for BFGS algorithm. 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 [1, +1]. So each variable is normalized in this range using the equation to improve the accuracy and efficiency of calculation:

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

where x_n is the normalized value of the corresponding x, $xmax$ 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 back-propagation. In Figure 9, 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 $(w_{kj}, k = 1, 2, \ldots, N)$ 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.

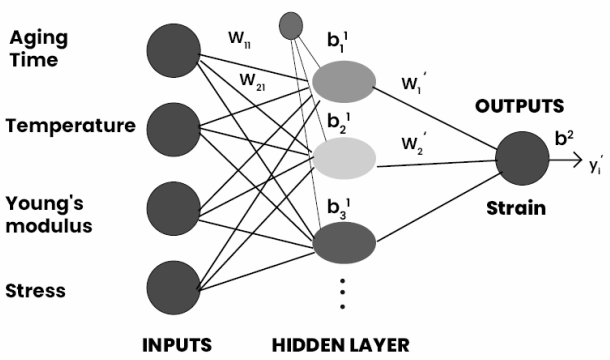

Fig. 9. Artificial neural network design with 4 inputs

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Step 1

For N neurons in hidden layer of the ANN and appropriate biases: $b_1^{(1)}, b_2^{(1)}...b_N^{(1)}$, 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:

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

Step 2
Now, the actived 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(.) 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:

$$\begin{array}{l} y'_i = g[\sum_{ki=1}^{N} w'_k f(b_k^{(1)} + \sum_{j=1}^{R} w_{kj}p_j) + b^{(2)}], \\ j = 1, 2...., R \quad k = 1, 2, ..., N. \end{array} \tag{3}$$

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

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 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 [21]; Dan. Foresee and Hagan 1997 [10]) 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:

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

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, 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:

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

where $E_w(w, M)$ is the sum of squares of architecture weights, M is the ANN architecture and α and β the regularization parameters or hyper-parameters. The second term on the right hand side of equation, E_w,is the $weight\ \ decay$ and with , the weight decay rate, favors small values of w and decreases the tendency of a model to overfit. Large values of α lead to posterior densities of weights that are highly concentrated around zero, so that the weights effectively disappear discounting connections in the network. If $\alpha << \beta$ then the training algorithm will make the errors smaller. If $\alpha >> \beta$, training will emphasize weight size reduction at the expense of network errors, thus producing a smoother network response [10].

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

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

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 α and β.

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:

$$\begin{array}{l} P(D|w, \beta, M) = (\frac{\beta}{\pi})^{n/2} \exp(-\beta E_D), \\ P(w|\alpha, M) = (\frac{\alpha}{\pi})^{N/2} \exp(-\alpha E_w), \end{array} \tag{7}$$

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:

$$\begin{array}{rl} P(w|D, \alpha, \beta, M) & = \frac{(\frac{\beta}{\pi})^{n/2}.(\frac{\alpha}{\pi})^{N/2} \exp(-(\beta E_D + \alpha E_w))}{P(D|\alpha, \beta, M)} \\ & = \frac{1}{Z_F(\alpha,\beta)}. \exp(-F(w)) \end{array} \tag{8}$$

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:

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

Now the equation 6 is, according to Mckay 1992 [21] :

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

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

where $Z_F(\alpha,\beta)$ can be estimated by Taylor series expansion: see (Foresee et al 1997 [10]). 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, α and β are calculated as:

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

with $\gamma = N - 2\alpha^{MP}Tr(\mathbf{H}^{MP})^{-1}$, 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} [10]. Mackay 1992 proposes an approach in [21]: 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

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

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

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

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 μ_{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 [32] 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.

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

We implement with this regularized function the BFGS algorithm.

This algorithm 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:

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

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.

REFERENCES

[1] 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. 10.1080/23311916.2020.1720894.

[2] 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 79:928–939.

[3] 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. 1-14. 10.1007/s00170-019-04486-4.

[4] H E. Balcıoglu, 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) 3267–3281. DOI: 10.1177/0021998315617998

[5] E. Burgaz, M. Yazici, M. Kapusuz, S. H. Alizir, H. Özcan. (2014). Prediction of thermal stability, crystallinity and thermomechanical properties of poly (ethylene oxide)/clay nanocomposites with artificial neural networks. Thermochimica Acta. 575. 159–166. 10.1016/j.tca.2013.10.032.

[6] 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):4479-4485.

[7] 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.

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

[8] 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):67-83.

[9] 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):15-32.

[10] F. Dan. Foresee and Martin T. Hagan. (1997) Gauss-Newton approximation to Bayesian learning. Proceedings of the International Joint Conference on Neural Networks, June.

[11] 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):453–463.

[12] ATC. Goh. (1995) Back-propagation neural networks for modeling complex systems. Artificial Intelligence in Engineering 9:143–15.

[13] 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, 353–384.

[14] Z. Zhang, P. Klein, and K. Friedrich. (2002) Dynamic mechanical properties of ptfe based short carbon fibre reinforced composites: experiment and artificial neural network prediction. Composites Science and Technology, 62(7-8):1001–1009.

[15] 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.

[16] 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):1497-1503, 2002.

[17] 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.

[18] S.Mahmoudi. (2017) Dynamique des structures composites linéaire et non-linéaire en présence d'endommagement. (Doctorial thesis). Université de Bourgogne Franche-Comté.

[19] https://www.mathworks.com/help/nnet/ref/trainbr.html

[20] 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, pages 115–133.

[21] D.J.C. Mackay. (1992) Bayesian interpolation. Neural Comput. 4, 415–447.

[22] I. Merdas, F. Thominette, A. Tcharkhtchi, and J. Verdu. (2002) Factors governing water absorption by composite matrices. Composites Science and Technology, 62(4):487-492.

[23] 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), 431- 441.

[24] M. Minsky S. P. Perceptrons. (1969). An Introduction to Computational Geometry Cambridge Ma.

[25] 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, 206–210.

[26] Y. Pei, K. Wang, M. Zhan, W. Xub, X. Ding. Thermal-oxidative aging of DGEBA/EPN/LMPA epoxy system : Chemical structure and thermalemechanical properties. Polymer Degradation and Stability, 96:1179-1186, 2011.

[27] 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):10733-10740.

[28] L.Poussines. (2012) Développement de nouveaux matériaux pour l'infusion de composites.(Doctorial thesis). ENIT Tarbes.

[29] 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, N°. 1, 1010–1021 https://doi.org/10.1080/14686996.2019.1673670

[30] F. Rosenblatt. (1962) Principles of neurodynamics.

[31] 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):355–365.

[32] 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.

[33] G.Z. Xiao and M.E.R. Shanahan. (1998) Swelling of DGEBA/DDA epoxy resin during hygrothermal ageing. Polymer, 39(14):3253-3260.

[34] J. Verdu. (2000) Action de l'eau sur les plastiques. Techniques de l'ingénieur, p.1-8.

[35] M. Wiciak-Pikuła, 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.

[36] 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.

ISBN: 978-988-14049-3-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)