Modeling the Temperature Dependence of Dynamic Mechanical Properties and Visco-Elastic Behavior of Thermoplastic Polyurethane Using Artificial Neural Network

Ivan Kopal 1,2,*, Marta Harničárová 1,3, Jan Valiček 1,3,4 and Milena Kušnerová 1,3

1 Institute of Physics, Faculty of Mining and Geology, Vysoká škola báňská—Technical University of Ostrava, 17. Listopadu 15, 708 33 Ostrava, Czech Republic; marta.harnicarova@vsb.cz (M.H.); jan.valicek@vsb.cz (J.V.); milena.kusnerova@vsb.cz (M.K.)
2 Department of Numerical Methods and Computing Modelling, Faculty of Industrial Technologies in Púchov, Alexander Dubček University of Trenčín, Ivana Krasku 491/30, 020 01 Puchov, Slovakia
3 Regional Materials Science and Technology Centre, Vysoká škola báňská—Technical University of Ostrava, 17. Listopadu 15, 708 33 Ostrava, Czech Republic
4 Technical Faculty, Slovak University of Agriculture in Nitra, Tr. A. Hlinku 2, 949 76 Nitra, Slovakia

* Correspondence: ivan.kopal@vsb.cz; Tel.: +420-597-323-559

Received: 11 September 2017; Accepted: 15 October 2017; Published: 18 October 2017

Abstract: This paper presents one of the soft computing methods, specifically the artificial neural network technique, that has been used to model the temperature dependence of dynamic mechanical properties and visco-elastic behavior of widely exploited thermoplastic polyurethane over the wide range of temperatures. It is very complex and commonly a highly non-linear problem with no easy analytical methods to predict them directly and accurately in practice. Variations of the storage modulus, loss modulus, and the damping factor with temperature were obtained from the dynamic mechanical analysis tests across transition temperatures at constant single frequency of dynamic mechanical loading. Based on dynamic mechanical analysis experiments, temperature dependent values of both dynamic moduli and damping factor were calculated by three models of well-trained multi-layer feed-forward back-propagation artificial neural network. The excellent agreement between the modeled and experimental data has been found over the entire investigated temperature interval, including all of the observed relaxation transitions. The multi-layer feed-forward back-propagation artificial neural network has been confirmed to be a very effective artificial intelligence tool for the modeling of dynamic mechanical properties and for the prediction of visco-elastic behavior of tested thermoplastic polyurethane in the whole temperature range of its service life.

Keywords: thermoplastic polyurethanes; visco-elastic properties; dynamic mechanical analysis; stiffness-temperature model; artificial neural networks

1. Introduction

Thermoplastic polyurethanes (TPUs) belong to the family of thermoplastic elastomers, which combine mechanical properties of rubber based materials with the good processability and recyclability of thermoplastics. They represent linear segmented semi-crystalline multi-block copolymers, composed of an alternation of short stiff hard crystalline segments and rather long soft amorphous flexible chains. The hard segments are made from di-isocyanate, while the soft segments consist of long and flexible polyether or polyester chains, which mutually connect two hard segments. In particular, the hard segments act as multi-functional tie points working as both...
polymers cross-links and reinforcing fillers. The domain structure formed by micro-phase separation of TPUs, due to the thermodynamic immiscibility or incompatibility between the hard and soft phase, endows these polymeric materials with their elastomeric properties, such that can be seen in cross-linked rubber networks [1,2]. The unique chemical structure of the TPUs results in a combination of interesting properties. They are characterized by a wide service temperature range, excellent tear and abrasion resistance, good resistance to non-polar solvents, high compression, tensile strength, and low-temperature elasticity. Therefore, TPUs are of great importance not only from an industrial but also from an academic point of view [3].

Similarly as of all polymeric materials, most of the TPUs properties are strongly dependent on temperature, which at different values induces different conformational changes in the polymer chains, leading to the corresponding transition processes that determine the behavior of polymers at various conditions [4]. In this sense, the influence of temperature on dynamic mechanical properties and visco-elastic behavior of TPUs plays a considerably important role and it is a longstanding subject of interest for many researchers. Nowadays, one of the most sensitive and most accurate methods used for study of dynamic mechanical properties, as well as of visco-elastic behavior of polymeric materials under various conditions (such as temperature, time, frequency of dynamic mechanical loading, and many others [5]), is represented by the dynamic mechanical analysis (DMA).

1.1. Dynamic Mechanical Analysis

DMA is a testing and analytical technique that is commonly used to characterize the mechanical responses of materials subjected to periodic dynamic stress in a cyclic manner, mostly the sinusoidal deformation force with very small amplitude. This technique helps engineers and scientists to understand and subsequently to predict the long-term material behavior under cyclic loading conditions. In the DMA test of polymeric material, a weak sinusoidal stress is applied at various frequencies and the strain inside a specimen of the known geometry is recorded as the temperature increases, which ultimately allows determination of the storage modulus $E'$, loss modulus $E''$, and the damping factor $\tan \delta$, as well as the identification of relaxation events that occurred in the tested polymer sample following the transition processes due to the conformational changes of its macromolecular chains. The storage modulus (stiffness) provides a measure of elastic energy stored in the material, the loss modulus (energy absorption or damping) refers to the amount of energy dissipated in the form of heat in each cycle of the sinusoidal deformation, while the ratio of the loss modulus to the storage modulus gives the damping factor. Due to the visco-elastic nature of the polymeric material, the sinusoidal strain lags the applied stress by a phase angle $\delta$, which is a measure of the visco-elastic damping or internal friction. The results obtained by investigations of dynamic mechanical properties or of visco-elastic behavior of polymeric materials using the DMA method are typically provided as a graphical representation of $E'$, $E''$ and $\tan \delta$ versus temperature or time at one or more frequencies of dynamic mechanical loading (Figure 1). It is then possible to identify in those diagrams all of the relaxation transitions that occurred in the tested polymer at corresponding transition temperatures [6].

1.2. Relaxation Transitions in Polymers

A key relaxation process in polymers is a glass transition at a temperature, $T_g$, which is frequently referred to as an $\alpha$-transition, $\alpha$-relaxation or as a primary relaxation. The $T_g$ or $T_\alpha$ point is defined as a temperature, at which the amorphous portion of polymeric materials experiences a physical change from a hard and brittle glass-like condition to very flexible rubbery-like state. Due to the fact that the glass transition event covers a range of temperatures without discontinuity in the measured, $T_g$, glass transition temperature is generally taken as a mid-point of this range. For amorphous portion of polymers, besides the glass transition, also one or more sub-$T_g$ phenomenon often referred to as secondary relaxations, can be observed. They are referred to as $\beta$, $\gamma$- and $\delta$-transitions at the temperatures $T_\beta$, $T_\gamma$ and $T_\delta$ and they appear in order of the descending temperature. These transition
processes are the result of local segmental motions that occur in a glassy state. By ‘local’, it is meant that only a small group of atoms are involved in the sub-$T_g$ processes, the pure existence of which proves that material in a glassy state is a dynamic one. The glass transition occurs when large segments of polymer chains start moving, $\beta$-transition is related to either localized movement in the main chain or the large side movement, and $\gamma$-transition relates with the movements occurring in the side chain of polymers. The $\delta$-transition is related to very small motions within macromolecule [7].

The authors Huh and Cooper in the work [8] have reported that TPUs may exhibit two separate glass transitions in the hard-and-soft-segment domains, respectively. The lower glass transition temperature $T_{\alpha ll}$ ($T_\gamma$) is associated with a change from stiff glassy behavior of the soft segments to their soft rubbery behavior and the higher one $T_{\alpha l}$ ($T_\alpha$) is due to the breakdown of hydrogen bonded interactions of Van der Waal’s forces between the rigid and flexible segments of TPUs.

Above the glass transition temperature, upon reaching the flow temperature $T_i$, the complete macromolecular chains can slip against each other. The amorphous structures of the polymer start to melt without chemical degradation of macromolecules. When the crystallite melting temperature $T_m$ is reached, the crystallites also start to melt. Due to the fact that the temperature range of crystallite melting exceeds the flow temperature of the amorphous state, the entire polymeric material will be plasticized. As long as no thermal degradation will occur in the molten phase at the decomposition temperature $T_d$, the polymer can reversibly get back into the solidified state by cooling and crystallite phases will again be generated. However, the size and distribution of crystallites can be now differing from the original status [9]. The crystalline and amorphous fraction of the semi-crystalline polymer below the $T_g$ temperature are present in the glassy state, within the temperature range between $T_g$ and $T_m$ the amorphous fraction will be in the rubbery state and above the temperature $T_m$ the crystalline fraction of the polymer will also already pass into the rubbery state, while the highly crystalline polymers do not have the rubbery area. The higher the amorphous fraction in the polymer, the stronger is its rubbery area. At a melting point of the crystalline phase, which is lower than the flow temperature of the amorphous phase $T_i$ and only in exceptional cases it is identical, the polymer suddenly passes from the rubbery state to the plastic state. The interval between the temperatures $T_m$ and $T_i$ depends on the molecular mass, the degree of crystallinity, as well as on several other parameters. In the case of a high share of the crystalline phase the importance of $T_g$ disappears because the polymer retains the properties of elastic body with the low deformation ability (glass) till the phase transition at $T_m$ temperature. However, the glass transition, melting of crystallites and thermal degradation of macromolecular chains are typical phase transitions virtually of all semi-crystalline polymers, including the TPUs [10].

![Figure 1](image-url). Temperature dependences of the average storage modulus $E'(T)$, loss modulus $E''(T)$, and damping factor tan $\delta(T)$ of tested thermoplastic polyurethanes (TPU) at a frequency of 1 Hz in the whole temperature range of its service life.
1.3. Stiffness-Temperature Model of Thermoplastic Urethanes

The authors Mahieux and Reifsnider in the work [11] showed, that the temperature dependence of the storage modulus $E'(T)$ at a given constant frequency of dynamic mechanical loading can be quantitatively described with high level of reliability by physically well substantiated unified analytical model based on a Weibull distribution of the breakage of secondary bonds between macromolecular chains during the primary and secondary relaxation processes in any polymeric systems in the form of

$$
E'(T) = \sum_{i=1}^{N} \Delta E'_i e^{-(\frac{T}{\Theta_i})^{m_i}},
$$

where $E'(T)$ is the temperature dependent storage modulus and $T$ is the absolute temperature of polymer, $\Delta E'_i$ are the storage modulus magnitudes of particular transition steps, the coefficients $\Theta_i$ represent absolute transition temperatures, the parameters $m_i$ are the Weibull moduli corresponding to the statistics of the secondary bond breakage, and $N$ is the number of relaxation events observed in the investigated polymeric material. This model has been transformed later into a robust physically based analytical model for prediction of the stiffness modulus for a wide range of temperatures and frequencies/strain rates by Richeton et al. in the study [12].

The stiffness-temperature model, presented by the Equation (1), in the form of exponential temperature function

$$
E'(T) = (E'_1 - E'_2) e^{-(\frac{T}{T_\alpha})^{m_\alpha}} + (E'_2 - E'_3) e^{-(\frac{T}{T_\beta})^{m_\beta}} + (E'_3 - E'_4) e^{-(\frac{T}{T_\gamma})^{m_\gamma}} + E'_4 e^{-(\frac{T}{T_f})^{m_f}},
$$

where $E'_1$ is the instantaneous storage modulus at the beginning of $E'(T)$ DMA curve, $E'_2$ is the instantaneous storage modulus immediately after the $\gamma$-transition, $E'_3$ is the instantaneous storage modulus immediately after the $\beta$-transition and $E'_4$ represents the instantaneous storage modulus at the beginning of the rubbery plateau (Figure 1), was successfully applied in one of our previous works in an analytical description of changes in the storage modulus of the tested TPU over a wide range of temperatures [13]. Both unknown physical parameters of the stiffness-temperature model (dynamic storage moduli and transition temperatures), as well as its statistical parameters (Weibull moduli) were estimated in the process of parametric fitting of the Equation (2) to the experimental $E'(T)$ DMA curve using a trust region algorithm for a non-linear least squares method. However, critical values of the storage modulus $E'_1 - E'_4$ can be determined from the $E'(T)$ DMA curve (Figure 1), the transition temperatures $T_\alpha$, $T_\beta$, $T_\gamma$ from the local maxima of the tan $\delta(T)$ curve (Figure 1) and the flow temperature $T_f$ from the onset of the constant value of $E'$ on the $E'(T)$ curve (Figure 1).

Accordingly, in the process of the multi-parametric fitting of experimental data set it is enough to estimate the unknown Weibull moduli $m_\alpha$, $m_\beta$, $m_\gamma$ a $m_f$, which quantify the width of the secondary bonds failure distribution corresponding to the level of crystallinit (m_α), the width of the secondary bond strengths distribution in the given temperature intervals (m_β, m_γ), and the width of secondary bond strength distribution corresponding to the TPU degree of cross-linking (m_f) [14]. At the same time, the temperature $T_{\text{all}}$ ($T_\gamma$) is associated with the glass transition of TPU soft segments, $T_{\text{all}}$ ($T_\alpha$) with the glass transition of its hard segments, $T_\beta$ can be attributed to a short-range order translation and to re-orientational motions within the PTU crystals, whereas $T_f$ corresponds to global translations of whole polymer molecules between entanglements [15]. In our work [13], it has been shown that the stiffness-temperature model presented by the Equation (2) is very successful in describing variations of the elastic behavior of the tested TPU with temperature from its glassy state up to the flow region and within the property transition regions. At the same time, this model makes it possible to predict the stiffness-temperature dependence of different types of TPUs, including the newly proposed ones, without necessarily having to subject them to time demanding and costly DMA testing.

Unfortunately, there is currently no physically justified unified analytical model that would be able to reliably describe also the viscous behavior of polymeric systems in the whole temperature range.
range throughout their service life, including all of the observed relaxation transitions in a way that Equation (1) describes their elastic behavior. However, the authors Ying Li et al. in work [16] present a predictive multi-scale computational framework to evaluate the visco-elastic properties of linear polymers based on their atomistic constituents. Many different multi-scale computational techniques are described in detail in [17]. Whereas, the artificial neural networks (ANNs) are, in general, excellent universal approximators not only of linear but also of non-linear functional dependencies between dependent and independent variables [18], it is quite natural to assume that temperature dependence of the loss modulus $E''(T)$ and damping factor $\tan \delta(T)$, as well as that of the storage modulus $E'(T)$, could be successfully modeled right by them.

1.4. Artificial Neural Networks

ANNs are intelligent mathematical computational methods inspired by biological neural networks and human brain that can learn from experience and generalize the learned things. Today they are used especially in the cases, where classic techniques fail, as well as other solutions for such problems, which can be formulated in the form of associative tasks [19]. ANNs are especially efficient in approximation of any continuous function with a finite number of discontinuities by mapping the relationships between vectors of the input (independent) and output (dependent) data [20]. With the right choice of structure and architecture they are capable to learn on the basis of the well-chosen input-output patterns to also model very complicated non-linear relations between the input and output data, for example to model a highly non-linear mechanical behavior of the materials such as polymers and their composites [21,22]. However, the acquired knowledge is, unfortunately, hidden in such ANN structures that cannot readily be extracted and interpreted. ANNs with multi-layer feed-forward architecture are the most commonly used neural networks in engineering applications thanks to their structural simplicity and relatively fast learning abilities [23].

The multi-layer feed-forward ANNs are non-parametric and robust mathematical models that are designed to find non-linear relationships between specified input-output data samples. This type of ANNs consists of at least of three layers of artificial neurons or processing units. The first layer is the input layer and the last layer is the output layer, while all of the layers between the input and output layers are referred to as the hidden layers. In this architecture of the ANNs, each neuron is connected to all of the neurons in the next layer but there are no connections between the neurons on the same layer. The signal propagates from the input layer, through all of the hidden layers, to the output layer in a forward direction. The numbers of neurons in the input, as well as in the output layers, are determined by the formulation of the investigated problem. The number of hidden layers and the number of neurons in these layers are related to the complexity of the relevant information contained in the learning or training data set. At the same time, the number of neurons in the hidden layers defines the capacity of the ANN. The knowledge about the relations between the input and output data is stored in the weights of connection between neurons, the values of which are updated during the supervised training process of the ANN with a set of known and representative values of the input-output data samples [24]. The training process starts with randomly set weights that are gradually updated by the minimization algorithm of squared difference between the target value and the computed output value on the output of each ANN neuron. Generally, to update the network weights iteratively, a technique called the back-propagation algorithm is used, which involves performing computations backward through the ANN [25]. The aim of the research described in the present work is exactly the modeling of the temperature dependencies of $E'(T)$, $E''(T)$, and $\tan \delta(T)$ of the investigated TPU submitted to DMA tests at a constant single frequency of dynamic mechanical loading in the whole temperature interval of its service life, or from its glassy state up to the flow region, using the multi-layer feed-forward ANNs utilizing the back-propagation learning algorithm.
2. Materials and Methods

Most of conventional TPU's are based either on polyester or polyether polyols, 4,4-diphenylmethane disiocyanate as an isocyanate component, and 1,4-butanediol as a chain extender [26]. In the present work, the synthesis of the tested TPU was performed in one liter of glass reactor by mixing two commercial reactive components—Axson Technologies PX 522/HT POLYOL and Axson Technologies PX 521-522 HT ISO ISOCYANATE (Axson Technologies, Newmarket Suffolk, UK)—under normal pressure and at room temperature. In order to ensure degassing and homogenization before polymerization, the reactive mixture was mixed for approx. 5 min with the use of an ultrasound mixer WELDER (WELDER, Hiroshima, Japan) in a VB vacuum chamber (VB, Prague, Czech Republic). Subsequently, the mixture was poured into a mold and left to cure at 80 °C for 4 h. The post-curing process proceeded at 100 °C for 16 h. Under these conditions the addition of catalyst was not necessary. From the resulting material, ten rectangular shaped samples with dimensions of 30 mm (length) × 6 mm (width) × 1.5 mm (thickness) were cut off using a Gravograph LS100 40 W CO2 laser cutter (Gravograph, Paris, France). The samples were then subjected to dynamic mechanical analysis over the entire temperature range of the TPU service life at a single constant frequency of dynamic mechanical loading.

The DMA test measurements were carried out using a Perkin Elmer Pyris Diamond 8000 Dynamic Mechanical Analyzer (Perkin Elmer, Baesweiler, Germany) in the uniaxial tensile mode at a constant frequency of dynamic mechanical loading of 1 Hz, with an amplitude of 20 µm from −125 up to 250 °C and at a constant heating rate of 3 °C·min⁻¹. The amplitude of applied dynamic stress was 0.1 MPa and the strain rate was 0.1 s⁻¹. The strain measurements were performed by a linear variable differential transformer in the DMA equipment with a span of 2 mm and a mean resolution of 2 nm. The measurements were performed on ten specimens cooled by liquid nitrogen from a Dewar flask (Perkin Elmer, Baesweiler, Germany). The average values of \( E', E'', \) and \( \tan \delta \) for each temperature \( T \) were computed. The measurement uncertainty was approx. 2%. The experimental results of DMA tests of the examined material in the form of temperature dependencies of the average storage modulus \( E'(T) \), loss modulus \( E''(T) \), and damping factor \( \tan \delta(T) \) at a single constant frequency of dynamic mechanical loading of 1 Hz in the whole temperature range of its service life are presented in Figure 1. The frequency of 1 Hz was chosen due to the fact that the secondary transitions and other structural features of the tested material could be easily detected, and the obtained results could be in the case of need more easily related to the experimental data acquired with the use of other techniques (e.g., by DSC method [27]). Figure 1 shows also highlighted critical values of the storage modulus \( E'_1 - E'_4 \) and temperatures of relaxation events \( T_{\alpha II}, T_{\beta}, T_{\alpha I} \) occurred in the tested TPU samples. In connection with the above described stiffness-temperature analysis of experimental data, the temperatures in the diagram in Figure 1 are presented in the absolute scale.

**Artificial Neural Network Modeling**

This work aimed at an investigation of temperature dependencies of the observed visco-elastic parameters \( E'(T), E''(T), \) and of \( \tan \delta(T) \) of the investigated TPU at a given constant frequency of its dynamic mechanical loading, three multi-layer feed-forward artificial neural network models were created utilizing the error back-propagation learning algorithm with an architecture consisting of one input, one hidden and one output layer [28]. The temperature history \( T(t) \) of TPU during the DMA test, representing a data set of the input experimental data, was naturally chosen as the sole input parameter of the ANN for all three of the created models, while the corresponding experimental values of the storage modulus, loss modulus, as well as damping factor, were chosen as a target ANN data set for each model individually. The input and output layer has one neuron, which corresponds with the number of the input and output variables of individual models, while it was necessary to optimize the number of neurons in a hidden layer with the use of a trial and error method for each individual model separately. The main criteria for this optimization were the maximum network performance and its trouble-free learning. A single hidden layer was chosen in view of the practical requirement for
the maximum network simplicity, its high stability, trouble-free learning, and minimal demands on computing power and computing time. Moreover, a number of mathematical works has shown that virtually any of the above described functions can be approximated with the required precision by means of a well-trained feed-forward ANN with a single hidden layer under the mild assumptions of the activation function (transfer function), which follows from the so-called universal approximation theorem for feed-forward ANNs [29].

Between the input and hidden layer of the ANN a hyperbolic tangent sigmoid transfer function was used in the form of [30]

\[ \text{tan sig}(n) = \frac{1 - e^{-2n}}{1 + e^{-2n}}, \]  

where \( n \) represents the net input into the hidden layer transfer function, which keeps the neurons output between certain limits as it is the case in the biological neuron. In order to avoid limitation of the network output by a small range of values, the pure linear transfer function

\[ \text{purelin}(m) = m, \]

where \( m \) is the net input into the transfer function of the output layer, was used between the hidden and output layer [30].

Generally, each neuron of the ANN represents its elementary building and computing unit located in the appropriate network layer. The input layer only transmits the input signal \( p \) to the hidden layer neurons without performing any calculations. Neurons of the hidden and output layer contain input, bias, weight, summing junction, transfer function, and output. At the input of each neuron of the hidden layer, each element \( p_i \) of the input data vector \( p \) with the length \( R^1 \) is multiplied by its weight \( w^{1,1} \). The weighted inputs of all \( S^1 \) neurons of the hidden layer are summed in their summing junctions. This weighted sum of the inputs together with the added bias \( b^1_j \) of the current neuron create an activation potential

\[ n^1_j = \sum_i w^{1,1} p_i + b^1_j \]  

leading to the activation of its output, while the bias is just a constant external input with a value of 1 multiplied by its weight, which represents the basic level of activity of the given neuron [31]. The activation signal, given by the Equation (5), or the net input of the hidden neuron \( n^1_j \) is then fed as the only argument into the non-linear transfer function, given by the Equation (3), which generates its output

\[ a^1_j = f^1(n^1_j) = \text{tan sig}(n^1_j) = \frac{1 - e^{-2n^1_j}}{1 + e^{-2n^1_j}}. \]

The total output of the hidden layer with \( S^1 \) hidden neurons, represented in a matrix form as

\[ a^1 = f^1(W^{1,1} p + b^1), \]

is then sent to the output layer.

Net input of the output layer

\[ n^2_j = \sum_j w^{2,1} a^1_j + b^2_j, \]

created by summing in the summing junction, is then inserted into the pure linear transfer function given by the Equation (4), which returns the output signal

\[ a^2_j = f^2(n^2_j) = \text{purelin}(n^2_j) = n^2_j, \]
or the output signal represented in a matrix form

\[ a^2 = f^2 (LW^{2,1}a^1 + b^2). \]  

(10)

Total output signal of the network \( y \) can be then presented as a matrix in the form

\[ y = a^2 = f^2 \left( LW^{2,1} \left( IW^{1,1}p + b^1 \right) + b^2 \right). \]  

(11)

The weight matrix \( IW^{1,1} \) represents a matrix of weights between neurons of the hidden and input layer, while \( LW^{2,1} \) represents the weight matrix between neurons of the hidden and output network layers. The upper index in the designation of all of the network elements represents its respective layer, it means that index 1 denotes a hidden layer and index 2 denotes an output layer. The lower index \( j \) belongs to the elements of the input vector \( p \), while the lower index \( i \) corresponds to the individual neurons in the respective layer (indicated by the upper index).

Due to the fact that the output of the hidden layer \( a^1 \) is also the input of the output layer with the output \( a^2 \), the hidden layer can be, in general, analyzed also as a single layer network with \( S^1 \) inputs, \( S^2 \) neurons and \( S^2 \times S^1 \) weight matrix

\[ W^2 = \begin{bmatrix} w^2_{1,1} & w^2_{1,2} & \cdots & w^2_{1,S^1} \\ w^2_{2,1} & w^2_{2,2} & \cdots & w^2_{2,S^1} \\ \vdots & \vdots & \ddots & \vdots \\ w^2_{S^2,1} & w^2_{S^2,2} & \cdots & w^2_{S^2,S^1} \end{bmatrix}. \]  

(12)

with weights \( w^2_{i,j} \), where the lower index \( q \) denotes an element of the input vector \( a^1 \), while the lower index \( r \) denotes the appropriate neuron, into which the signal from this element is brought [32]. Each neuron of the hidden layer of the proposed ANN is connected by means of weights with each element of the input vector and with neuron of the output layer, and by bias also with external environment of the network.

The ANN architecture for modeling the temperature dependencies of all three observed visco-elastic parameters is in an abbreviated notation [32], schematically shown in Figure 2. The input vector \( p \) in our case for all generated ANN models represents a vector of the temperature history \( T \) of the TPU during its DMA test, while the output signal \( y \) of the network for individual models are the vectors of the corresponding values \( E'(T), E''(T), \) or tan \( \delta(T) \) at temperatures, \( T(i) \) predicted by the network. Vectors and matrices are in this work marked in straight bold fonts.

\[ \begin{array}{c}
\text{Input Layer} \\
\begin{array}{c}
T \\
R \times 1
\end{array}
\end{array} \quad \begin{array}{c}
\text{Hidden Layer} \\
\begin{array}{c}
\text{IW}^{1,1} \\
S^1 \times R \\
\begin{array}{c}
b^1 \\
S^1 \times 1
\end{array}
\end{array}
\end{array} \quad \begin{array}{c}
\text{Output Layer} \\
\begin{array}{c}
\text{LW}^{2,3} \\
S^2 \times S^1 \\
\begin{array}{c}
b^2 \\
S^2 \times 1
\end{array}
\end{array}
\end{array} \quad \begin{array}{c}
\text{\( y = a^2 \)} \\
S^1 \times 1
\end{array}
\begin{array}{c}
\text{\( E' \)} \\
\text{\( E'' \)} \\
\text{\( \tan \delta \)}
\end{array}
\]

**Figure 2.** Schematic of the artificial neural networks (ANN) architecture for modeling the temperature dependencies of the storage modulus \( E'(T) \), loss modulus \( E''(T) \) and damping factor \( \tan \delta(T) \) of the investigated TPU at a given constant frequency of its dynamic mechanical loading.

At first, it was necessary to teach the created ANN to recognize the relationships between the input and output data of the DMA test, then to verify the reliability of the learned recognition and finally to test its ability to generalize the acquired and verified knowledge at prediction of new experimental results of the DMA test. The input and target data of the created ANN were therefore randomly
divided to training, validation, and testing data sets in the proportion 0.70:0.15:0.15, respectively. There is no principle limitation for these divisions, but in generally, it is common that the training data set has larger amounts of data than the other sets [33]. On the other hand, the number of training data can strongly influence the ANN predictive quality [34].

The training data were used for supervised learning of the network [35], during which the difference between the values $y_k$ of the output vector $y$ predicted by the network and the values $d_k$ of the vector of the desired (training) data $d$ is minimized with the use of step by step modification of the weights $IW^{1,1}$, $IW^{2,1}$ and biases $b^1$, $b^2$. Values of the error vector

$$\delta = d - y$$  \hspace{1cm} (13)

is appreciated with the use of the mean squared error (MSE) function

$$E(w, b) = \frac{1}{2l} \sum_{k=1}^{l} (d_k - y_k)^2 , \hspace{1cm} (14)$$

where $l$ is the total number of its elements, which is minimized during the ANN training process [36], which leads to an adaptation of weights and biases of the network to the training data, or to its optimization of the training data set.

The Levenberg–Marquardt back-propagation minimization algorithm [37], also known as the damped least squares method, based on the modification of gradient Gauss–Newton method [38], was used for training of all three ANN models. The iterative scheme of the Levenberg–Marquardt algorithm for the modification of weights and biases with the error function MSE has the form:

$$z_{k+1} - z_k = \left( J_k^T J_k + \lambda_k I \right)^{-1} J_k^T E_k , \hspace{1cm} (15)$$

where $z$ represents the weights $w$ and biases $b$ of the network, the product $J_k^T J_k$ is referred to as Hessian matrix approximation for non-linear least square optimization [39], $J_k$ is the Jacobian matrix that contains first derivatives of the network MSE with respect to the weights and biases [40], $I$ is the identity matrix [41], $\lambda_k$ is the non-negative learning parameter [42] and the symbol $^T \{ displaystyle \hat{\mathsf{T}} \}$ denotes the matrix transpose [43]. The iterative learning process starts with a random adjustment of the weight and its adaptation continues until reaching the specified maximum number of iterations (epochs), or until the MSE network is minimized to a pre-determined target value. The learning process, during which the ANN seeks through the step-by-step modification of weights and biases the most accurate training data model, must be terminated before the network also begins to adapt to the noise present in each experimental data, which causes a dramatic reduction of its predictive abilities for the unknown data. Therefore, during minimization of the error function, a development of the error function in the validation data set is also monitored. In the event that error of the validation data begins to increase during several iterative cycles, while the error of the training data is still decreasing or is no longer changing, the training process is prematurely terminated in order to prevent the so called overfitting of the network, the cause of which is usually too high a number of neurons in the hidden layer [44]. The weights and biases, at which the MSE of the validation data set was minimal, are then taken as the resulting values. It means that the training data set is used to fit the model by finding its unknown but correct weights and biases, whereas the validation data set is used to test how well the model performs on the unseen data, or for debugging the ANN architecture (the number of hidden neurons or possibly also hidden layers).

The testing dataset contains the data, with which the ANN did not come into contact during the learning process or during validation, and which serve for verification of the level of network learning or for verification of its ability to generalize the acquired knowledge stored in the values of weights and biases values at prediction of new experimental results of the DMA test at new temperatures, unknown to the network. A significant difference in the number of iterations needed to achieve a
minimum MSE of validation and testing data may indicate an inappropriate ratio of the division of
the input-output data of the network to the training, validation, and testing data. The inability
of the network to describe more significant non-linearities in the testing data set, or the so called ANN
underfitting, is usually caused by an insufficient number of hidden neurons or hidden layers [44].

The ANN adaptation begins with random values of weights, that is why it is quite natural that
each individual solution found in the process of its optimization may be completely different from
any other, and in general, it is not possible to guarantee that the solution found is the best possible
one. Even very different setting of the ANN weights can provide a practically identical prediction for
training data, but the difference may occur in quality of the prediction for the unknown data. It is
therefore always reasonable to repeat the training, validation, and testing of the ANN several times
and to choose the model with the smallest MSE of the training, validation, and testing data.

The only reliable way to find an optimal number of neurons in the hidden ANN layer is the trial
and error method. In this work, the initial values $N_h$, calculated according to the empirical relation
below, were used for an effective search of an optimal number of hidden neurons

$$N_h = \frac{N_s}{\alpha(N_i + N_o)}, \quad (16)$$

where $N_s$ is number of the samples in the training data set, $N_i$ number of the input neurons, $N_o$ number
of the output neurons, and $\alpha$ is a scaling factor with a value between 2–10, expressing the requirement
of how general the ANN model should be, or to what extent the possibility of the network overfitting
must be suppressed. In this work, the network training, validation and testing process for each selected
$N_h$ was repeated 50 times. As the optimal number of neurons in the hidden layer was taken the one,
at which the MSE for all three data selections was minimal and the value of the linear correlation
coefficient $R$ between the output and the desired data

$$R = \frac{N_s}{\alpha(N_i + N_o)}, \quad (17)$$

approached the value of one in the closest possible manner. The covariance of the output and
desired data is marked by the symbol $\text{Cov}(d,y)$, while the symbols $s_d$ and $s_y$ represent their standard
deviations, respectively [45].

The design of the structure and optimization of the ANN architecture for modeling of temperature
dependencies of the storage modulus, loss modulus, and damping factor of the investigated TPU
were performed in the development environment “Neural Network Toolbox” of the software package
Matlab® Version 7.10.0.499 R2010a 64-bit (MathWorks, Natick, MA, USA), which provided a complete
set of tools for sufficiently efficient work with artificial neuron networks of various types [32]. As it
was already mentioned above, the Levenberg-Marquardt algorithm was used for training the network,
the input-target data of which were randomly divided to training, validation and testing data sets
depending on the ratio given above. The various parameter settings for the ANN training are the
maximum number of epochs to train: 1000, minimum performance gradient: $1^{-10}$, initial value for
learning parameter: $1^{10}$, and the maximum validation failures: 6.

3. Results

Figure 1 shows the results obtained from DMA tests of the studied material. The $E'(T)$, $E''(T)$
and $\tan \delta(T)$ curves exhibit a typical shape for the high molecular weight, weakly semi-crystalline,
thermoelastic elastomers throughout the temperature range below the flow region. The $\alpha$-, $\beta$-
and $\gamma$-transitions occur as a direct result of the three phase morphology, in particular crystalline
hard segments phase, crystalline soft segments phase, and amorphous soft segments phase of TPU.
The temperature $T_{\alpha I}$ corresponds to the glass transition of soft segments, $T_{\alpha II}$ corresponds to the
glass transition of hard segments, and $T_\beta$ corresponds to the short-range order translation and
reorientational motions within the crystals [8]. A sharp drop of $E'(T)$ at temperatures higher than $T_{\alpha II}$
indicates a significant phase separation induced by the symmetrical structure of rigid segments, a low crosslinking level and a small degree of crystallinity of TPU. The slope of the glass transition of the rigid segments, observed on the curve \(\tan \delta(T)\), shows the homogeneous structure of the physical network and suggests a relatively narrow variance of the molecular weights.

The temperature independent rubber plateau corresponds to uniform rigid segments forming stable crystals, which soften in a relatively narrow temperature range, as evidenced by the increase in \(\tan \delta(T)\) at the end of the DMA curve. It was proved in [13], that the stiffness-temperature model (2) is able to describe only the functional dependence \(E'(T)\) with a high degree of reliability, i.e., the elastic behavior of the examined TPU.

The ANN, the schematic of architecture of which is presented in Figure 2, provides for \(E'(T)\) the results with the lowest MSE and the highest accuracy \(R\) for all three data sets (training, validation and testing) simultaneously at 14 neurons, for \(E''(T)\) at 13 neurons and for \(\tan \delta(T)\) at 15 neurons in a hidden layer. Taking into account that the input and output layer of all three ANN models contains a single neuron, the optimal architecture of the multi-layer feed-forward back-propagation ANN for modeling of temperature dependence of the storage modulus \(E'(T)\), loss modulus \(E''(T)\) and damping factor \(\tan \delta(T)\) of the investigated TPU at given constant frequency of its dynamic mechanical loading in the whole temperature range of its service life can be presented in numeric form as 1-14-1, 1-13-1 and 1-15-1, respectively [46].

The ANNs do not usually enable a calculation of statistical parameters and diagnostics for a detailed assessment of the model quality, with the use of for example variance of regression coefficients, F and t statistics for the testing the significance of the model, diagnostic graphs, etc., as it is in the case of parametric regression models of experimental data [47]. For this reason, it is necessary to use different methods for the assessment of quality of the ANN model [48]. In this work, in order to validate the ability and accuracy of the modeled results related to all of three investigated visco-elastic parameters, first of all of the performance plots of the ANN models have been constructed and they are shown in Figure 3. These plots show how the MSE function minimizes during the training process. After six validation checks the network stops its training and returns to the state, in which the minimum validation performance was observed. This is indicated by the green circle in each figure. It follows from the plot in Figure 3a that the best validation performance of the ANN model for \(E'(T)\) occurred at the epoch 163 with the associated MSE value of \(3.4304 \times 10^{-6}\), for \(E''(T)\) at the epoch 36 with MSE value of \(7.4298 \times 10^{-7}\) (Figure 3b) and for \(\tan \delta(T)\) at the epoch 274 with MSE value of \(2.0137 \times 10^{-7}\) (Figure 3c). The final MSE of all three ANN models is rather negligible, the training data set MSE, the validation data set MSE, as well as the test data set MSE have similar characteristics. No significant overfitting has occurred by iteration (where the best validation performance has occurred), which has indicated the achievement of excellent results for both the approximation and the prediction of values of the investigated functions [49]. However, the training of the ANN model for \(E''(T)\) is the least time-consuming, as this model converges to the minimal MSE value already in the epoch 36 of the training process, while the most time-consuming training of the ANN model for \(\tan \delta(T)\) converges to the MSE minimum only in the epoch 274.
The MSE function is a performance measurement adopted to determine the ANN performance, while $R$ values measure the correlation between the outputs and targets. In Figure 4, the linear regression plots of targets relative to outputs or fitness plots for all three data sets of the ANN (training, validation, and testing) and entire data used in the network models for $E'(T)$ (Figure 4a), $E''(T)$ (Figure 4b), and $\tan \delta(T)$ (Figure 4c) are presented. In the top and left part of these figures the $R$ values and equations of linear fit lines in the format

$$Output \sim slope \cdot Target + \text{intercept}$$

are shown, respectively.

It is apparent from the linear regression plots presented in Figure 4, which practically all of the modeled data (hollow circles) fall on a 45° line expressing the exact agreement between the computed values (output data) and experimental ones (target data) with the associated $R$ value of 1, which confirms the achievement of excellent results of all three analyzed ANN models [44]. Lower values of the correlation coefficients $R$ in the case of $E''(T)$ of the ANN model in comparison with other two models document, that the nonlinear behavior of TPU damping under the given conditions is higher than that of its elastic resistance or stiffness. Namely, the different level of non-linearity of the temperature dependencies of individual monitored visco-elastic parameters is the reason why we did not propose for solution of our problem a single ANN model with three outputs but a single-output ANN model for each individual parameter separately.
The difference between the experimental and modeled values of $E'(T)$, $E''(T)$, and $\tan \delta(T)$ are presented in Figure 5, in detail. It can be seen that error of all three ANN models on the level of the order of $10^{-3}$ is negligibly small in comparison with the experimental data acquired from the DMA tests.

After being sure about the accuracy and authenticity of the results, for any unknown value of the temperature from the whole investigated temperature range, the corresponding value of the storage modulus, loss modulus, as well as damping factor of the examined TPU can be obtained by simulation of the above described ANN models using the found numerical amounts of weights and biases for each layer [50].
The ANN is used by many researchers for the modeling of dynamic mechanical properties and viscoelastic behavior of TPUs, which are currently among the most used construction materials in many practical applications.

The elastic behavior of composites with desired dynamic–static properties, the development of new TPUs, and the prediction of their viscoelastic behavior over the wide range of temperatures are important areas of research. The artificial neural network technique has been used to model the temperature dependence of dynamic mechanical properties of the thermoplastic polyurethane within the whole temperature range of its service life.

After being sure about the accuracy and authenticity of the results, for any unknown value of the temperature from the whole investigated temperature range, the difference between the experimental and modeled values of the storage modulus, loss modulus, as well as damping factor of the examined TPU can be obtained by biases for each layer.

Figure 5. (a) Errors of the ANN model for $E'(T)$; (b) Errors of the ANN model for $E''(T)$; and, (c) Errors of the ANN model for $\tan \delta(T)$.

The excellent predictive abilities of ANN models can also be used in the analysis of viscoelastic properties of polymers. The multi-layer feed-forward back propagation artificial neural network, utilizing the Levenberg–Marquardt learning algorithm, has been used. The hyperbolic tangent sigmoid transfer function between the input and hidden layers with one hidden layer is used. The excellent agreement between the modeled and experimental data has been found over the entire investigated temperature interval, including all observed relaxation transitions. The multi-layer analysis of the elastic behaviors of the tested thermoplastic polyurethane and for prediction of the viscoelastic properties of the composite materials and polymer based materials, including all observed relaxation transitions, is possible.

In the presented work, the artificial neural network technique has been used to model the dynamic mechanical properties and viscoelastic behavior of TPUs, which are currently among the most used construction materials in many practical applications.

The authors declare no conflict of interest.

Conflicts of Interest: The authors declare no conflict of interest.
4. Conclusions

In the presented work, the artificial neural network technique has been used to model the temperature dependence of dynamic mechanical properties and viscoelastic behavior of thermoplastic polyurethane over the wide range of temperatures. The variations of the storage modulus, loss modulus and damping factor with temperature, obtained from the dynamic mechanical analysis tests across transition temperatures at constant single frequency of its dynamic mechanical loading, have been modeled by three models of well-trained, multi-layer feed-forward back-propagation artificial neural network, utilizing the Levenberg–Marquardt learning algorithm, with one hidden layer, the hyperbolic tangent sigmoid transfer function between the input and hidden layers, and the pure linear function between the hidden and output layers. The excellent agreement between the modeled and experimental data has been found over the entire investigated temperature interval, including all of the observed relaxation transitions. The multi-layer feed-forward back-propagation artificial neural network has been confirmed to be a very effective artificial intelligence tool for the modeling of dynamic mechanical properties and for prediction of viscoelastic behavior of the tested thermoplastic polyurethane within the whole temperature range of its service life. The ANN is used by many researchers for the analysis of viscoelastic properties of polymers. There is no work up to date dealing with ANN modeling of dynamic mechanical properties and the viscoelastic behavior of TPUs, which are currently among the most used construction materials in many practical applications. The excellent predictive abilities of ANN models can also be used in the development of new TPUs-based materials and polymer composites with desired dynamic mechanical properties, relaxation transitions, and predictable viscoelastic behaviors.

Acknowledgments: This research work has been supported by the Slovak grant agency KEGA the project No. 005TnUAD-4/2016 and the project SGS SP2017/116.

Author Contributions: Ivan Kopal wrote the paper and designed the experiments; Marta Harničárová performed the experiments; Jan Valiček and Milena Kušnerová verified the model.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Drobny, J.G. Handbook of Thermoplastic Elastomers, 2nd ed.; Elsevier: Oxford, UK, 2014.
2. Krmela, J.; Krmelová, V. Dynamic Experiment of Parts of Car Tyres. Procedia Eng. 2017, 187, 763–768. [CrossRef]
3. Patton, S.T.; Chen, C.; Hu, J.; Grazulis, L.; Schrand, A.M.; Roy, A.K. Characterization of Thermoplastic Polyurethane (TPU) and Ag-Carbon Black TPU Nanocomposite for Potential Application in Additive Manufacturing. Polymers 2017, 9, 6. [CrossRef]
4. Prisacariu, C. Polyurethane Elastomers: From Morphology to Mechanical Aspects; Springer: Wien, Austria, 2011.
5. Ward, I.M.; Sweeney, J. Mechanical Properties of Solid Polymers, 3rd ed.; Wiley: Hoboken, NJ, USA, 2012.
6. Ramachandran, V.S.; Paroli, R.M.; Beaudoin, J.J.; Delgado, A.H. Handbook of Thermal Analysis of Construction Materials; Noyes Publications/William Andrew Pub: Norwich, NY, USA, 2002.
7. Menczel, J.D.; Prime, R.B. Thermal Analysis of Polymers: Fundamentals and Applications; Wiley: Hoboken, NJ, USA, 2008.
8. Huh, D.S.; Cooper, S.L. Dynamic mechanical properties of polyurethane block polymers. Polym. Eng. Sci. 1971, 11, 369–376. [CrossRef]
9. Gubbott, P. Principles and Applications of Thermal Analysis; Wiley: Hoboken, NJ, USA, 2007.
10. Ward, I.M.; Sweeney, J. An Introduction to the Mechanical Properties of Solid Polymers, 2nd ed.; Wiley: Chichester, UK, 2004.
11. Mahieux, C.A.; Reifsnyder, K.L. Property modeling across transition temperatures in polymers: A robust stiffness-temperature model. Polymer 2001, 42, 3281–3291. [CrossRef]
12. Richeton, J.; Schlatter, G.; Vecchio, K.S.; Rémont, Y.; Ahzi, S. Unified model for stiffness modulus of amorphous polymers across transition temperatures and strain rates. Polymer 2005, 46, 8194–8201. [CrossRef]
13. Kopal, I.; Bakošová, D.; Koštial, P.; Jančíková, Z.; Valiček, J.; Harničárová, M. Weibull distribution application on temperature dependence of polyurethane storage modulus. *Int. J. Mater. Res.* 2016, 107, 472–476. [CrossRef]

14. Rinne, H. *The Weibull Distribution: A Handbook*; CRC Press: Boca Raton, FL, USA, 2009.

15. Brinson, H.F.; Brinson, L.C. *Polymer Engineering Science and Viscoelasticity*, 2nd ed.; Springer: Berlin, Germany, 2014.

16. Li, Y.; Tang, S.; Abberton, B.C.; Kröger, M.; Burkhart, C.; Jiang, B.; Papakonstantopoulos, G.J.; Poldreiff, M.; Liu, W.K. A predictive multiscale computational framework for viscoelastic properties of linear polymers. *Polymer* 2012, 53, 5935–5952. [CrossRef]

17. Li, Y.; Abberton, B.C.; Kröger, M.; Liu, W.K. Challenges in multiscale modeling of polymer dynamics. *Polymers* 2013, 5, 751–832. [CrossRef]

18. Fausett, L.V. *Fundamentals of Neural Networks*, 1st ed.; Prentice Hall: New York, NY, USA, 1994.

19. Zhang, Z.; Friedrich, K. Artificial neural networks applied to polymer composites: A review. *Compos. Sci. Technol.* 2003, 63, 2029–2044. [CrossRef]

20. Sifaoui, A.; Abdelkrim, A.; Benrejeb, M. On the Use of Neural Network as a Universal Approximator. *Int. J. Sci. Tech. Control Comput. Eng.* 2008, 2, 386–399.

21. Bessa, M.A.; Bostanabad, R.; Liu, Z.; Hu, A.; Apley, D.W.; Brinson, C.; Chen, W.; Liu, W.K. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. *Comput. Methods Appl. Mech. Eng.* 2017, 320, 633–667. [CrossRef]

22. Le, B.A.; Yvonnet, J.; He, Q.C. Computational homogenization of nonlinear elastic materials using neural networks. *Int. J. Numer. Meth. Eng.* 2015, 104, 1061–1084. [CrossRef]

23. Norgaard, M.; Ravn, O.; Poulsen, N.K.; Hansen, L.K. *Neural Networks for Modeling and Control of Dynamic Systems*; Springer: Berlin, Germany, 2002.

24. Jin, D.; Lin, A. *Advances in Computer Science, Intelligent System and Environment*; Springer: Berlin, Germany, 2011.

25. Aliiev, R.; Bonfig, K.; Aliiev, F. *Soft Computing*; Verlag Technik: Berlin, Germany, 2000.

26. Bhowmick, A.K.; Stephens, H.L. *Handbook of Elastomers*, 2nd ed.; CRC-Press: Boca Raton, FL, USA, 2000.

27. Qi, Z.N.; Wan, Z.H.; Chen, Y.P. A comparative DSC method for physical aging measurement of polymers. *Polym. Test.* 1993, 12, 185–192. [CrossRef]

28. Freeman, J.A.; Skapura, D.M. *Neural Networks Algorithms, Applications, and Programming Techniques*; Addison-Wesley Publishing Company: Boston, MA, USA, 1991.

29. Auer, P.; Burgsteiner, H.; Maass, W.A. A learning rule for very simple universal approximators consisting of a single layer of perceptrons. *Neural Netw.* 2008, 21, 786–795. [CrossRef] [PubMed]

30. Beale, M.H.; Hagan, M.T.; Demuth, H.B. *Neural Network Toolbox™ User’s Guide*; The MathWorks, Inc.: Natick, MA, USA, 2017.

31. Bowden, G.J.; Maier, H.R.; Dandy, G.C. Optimal division of data for neural network models in water resources applications. *Water Resour. Res.* 2002, 38, 2-1–2-11. [CrossRef]

32. Trebar, M.; Susteric, Z.; Lotric, U. Predicting mechanical properties of elastomers with neural networks. *Polymer* 2007, 48, 5340–5347. [CrossRef]

33. Mohri, M.; Rostamizadeh, A.; Talwalkar, A. *Foundations of Machine Learning*; MIT Press: Cambridge, MA, USA, 2012.

34. Hagan, M.T.; Menhaj, M. Training feed-forward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* 1999, 5, 989–999. [CrossRef] [PubMed]

35. Croeze, A.; Pittman, L.; Reynolds, W. *Nonlinear Least-Squares Problems with the Gauss-Newton and Levenberg-Marquardt Methods*; University of Mississippi: Oxford, UK, 2012.

36. Fairbank, M.; Alonso, E. Efficient Calculation of the Gauss-Newton Approximation of the Hessian Matrix in Neural Networks. *Neural Comput.* 2012, 24, 607–610. [CrossRef] [PubMed]

37. Wackerly, D.; Mendenhall, W.; Scheaffer, R.L. *Mathematical Statistics with Applications*, 7th ed.; Thomson Higher Education: Belmont, TN, USA, 2008.
41. Barnes, R.J. Matrix Differentiation; University of Minnesota: Minneapolis, MN, USA, 2017.
42. Madić, M.J.; Radovanović, M.R. Optimal Selection of ANN Training and Architectural Parameters Using Taguchi Method: A Case Study. FME Trans. 2011, 39, 79–86.
43. Nocedal, J.; Wright, S.J. Numerical Optimization, 2nd ed.; Springer: Berlin, Germany, 2006.
44. Zhang, T.; Yu, B. Boosting with Early Stopping: Convergence and Consistency. Ann. Stat. 2005, 33, 1538–1579. [CrossRef]
45. Syed, A.H.; Munir, A. Distribution of Mean of Correlation Coefficients: Mean of Correlation Coefficients; LAP Lambert Academic Publishing: Saarbrücken, Germany, 2012.
46. Nematollahi, M.; Jalali-Arani, A.; Golzar, K. Organoclay maleated natural rubber nanocomposite. Prediction of abrasion and mechanical properties by artificial neural network and adaptive neuro-fuzzy inference. Appl. Clay Sci. 2014, 97–98, 187–199. [CrossRef]
47. Simon, L. An Introduction to Multivariable Mathematics; Morgan & Claypool: St. Louis, CA, USA, 2008.
48. Shi, F.; Wang, X.C.; Yu, L.; Li, Y. MATLAB 30 Case Analysis of MATLAB Neural Network; Beijing University Press: Beijing, China, 2009.
49. Wong, S.S.M. Computational Methods in Physics and Engineering, 2nd ed.; World Scientific Publishing: New York, NY, USA, 1977.
50. Suzuki, K. Artificial Neural Networks—Industrial and Control Engineering Applications; InTech: Rijeka, Croatia, 2011.

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).