Mechanical Property Classification of Vapor-Grown Carbon Nanofiber/Vinyl Ester Nanocomposites Using Support Vector Machines

O. AbuOmar[1,2], S. Nouranian[2], R. King[1,2], T.M. Ricks[3], T.E. Lacy[3]
[1]Department of Electrical and Computer Engineering, Mississippi State University, MS 39762, USA
[2]Center for Advanced Vehicular Systems (CAVS), Mississippi State University, MS 39762, USA
[3]Department of Aerospace Engineering, Mississippi State University, MS 39762, USA

Abstract — *In the context of materials informatics, the support vector machines technique was used to analyze and classify a large dataset of vapor-grown carbon nanofiber (VGCNF)/vinyl ester (VE) nanocomposites into three classes of desired mechanical properties, i.e., high storage modulus, high true ultimate strength, and high flexural modulus. Resubstitution and 3-folds repetitive cross validation techniques were implemented and the resulting classification information was compared and analyzed through sets of confusion matrices. This classification proves to be useful to materials designers and engineers, since a qualitative assessment of the expected nanocomposite mechanical response is given when suitable changes are made to the formulation, processing, and environmental conditions. This classification accelerates the lead time for the development of VGCNF/VE nanocomposites for a specific engineering application.*

Keywords: Support vector machines, materials informatics, vapor-grown carbon nanofiber/vinyl ester, confusion matrix, repetitive cross validation.

1 Introduction

The support vector machine (SVMs) technique [1] is considered one of the most widely used techniques in artificial intelligence community. This technique employs datasets of different sizes and dimensions and from different fields and domains. SVMs can be used for both supervised and unsupervised learning problems. Unsupervised learning ideally requires a large number of data vectors (points) within a particular dataset in order to adequately model a problem and avoid over-training (over-fitting). Supervised learning, however, can be utilized with a less number of data vectors, but some prior knowledge of the problem is needed in order to assist the SVMs model in generalizing and predicting the correct quantity given an unknown data vector [1]. SVMs can also assign linearly and nonlinearly separable data into two or more classes [1]. SVMs have recently been employed in the context of materials science and engineering to extend materials informatics [2-5]. This interdisciplinary study integrates computer and information science with other knowledge domains to facilitate knowledge discovery. For example, materials scientists can use materials informatics to interpret acquired experimental data through the use of SVMs and other machine learning approaches. It can also accelerate the research process and guide the development of new materials with desired mechanical properties. Materials informatics is being fueled by new and dynamic growth in the information technology sector and is driving the interest in SVMs, data mining, machine learning, information retrieval, and other knowledge representation or discovery schemes in the engineering disciplines [6]. AbuOmar, *et al*. [7] applied an artificial neural network (ANN) technique to a dataset associated with the viscoelastic response of a vapor-grown carbon nanofiber (VGCNF)/vinyl ester (VE) nanocomposite material system. The ANN was trained using the resubstitution method and the 3-fold repetitive cross validation (RCV) technique to provide a predictive model for these responses with minimal mean square error [7]. Roberts, *et al*. [8] presented a model that classifies different materials based on their microstructure. The core of the designed model is an SVMs classifier that identifies the appropriate class of given material sample based on microstructural characteristics such as Haralick variables, the Euler parameter, and the fractal dimension [8]. Swaddiwudhipong, *et al*. [9] utilized another important and efficient materials informatics technique, *i.e*., least squares support vector machines (LS-SVMs) [10]. Four LS-SVMs models simulated the relationship between the indentation load-displacement characteristics and elasto-plastic material properties, which are subject to the law of power hardening. No iterative approaches were used and it was found out that LS-SVMs approach was robust in determining the parameters in this relationship. Hu, *et al*. [11] used materials informatics to resolve the problem of materials science image data sharing. An ontology-based approach was employed to develop annotation for non-structured materials science data with the aid of semantic web technologies. Sabin, *et al*. [12] evaluated an alternative statistical Gaussian process model, which assumes a normal probability distribution over all of the training data and then interpolates to make predictions of microstructural evolution arising from static recrystallization in a non-uniform strain field. In this work, a specific class of advanced engineering materials was studied i.e., polymer nanocomposites [13]. These materials have multifunctional properties and are increasingly being used for aerospace, automotive, biomedical, fuel cell, catalysis, and other applications. For example, improved stiffness properties and energy absorption characteristics are desired in automotive structural applications and nano-enhanced polymer composites meet these requirements [14]. They have been the subject of intensive research in recent years [15, 16]. AbuOmar, *et al*. [17] applied data mining and knowledge discovery techniques to a thermosetting VGCNF/VE nanocomposite material system [18-21] and include self-organizing maps (SOMs) [22, 23] and clustering techniques [24, 25]. The SOMs were used to recommend VGCNF/VE nanocomposite systems exhibiting the same storage and loss modulus responses in order to minimize the material preparation cost. A clustering technique (*i.e*., a fuzzy C-means algorithm) was also applied to discover any pattern in the nanocomposite behavior after using principal component analysis (PCA) as a dimensionality reduction technique [26].

This study seeks to expand the current knowledge of the influence of formulation, processing, and environmental factors on the mechanical behavior of VGCNF/VE nanocomposites by including a wider range of measured mechanical properties, *i.e*., viscoelastic property data [18], compressive and tensile property data, and flexural property data [27]. This new dataset provides a more general insight into the mechanical behavior of VGCNF/VE nanocomposites for data mining purposes. Application of data mining and knowledge discovery techniques to a comprehensive dataset of mechanical responses of polymer nanocomposites is unprecedented and novel. In the context of materials informatics, the results of this study serve as a guideline for materials scientists and engineers to efficiently design or optimize a material system for a certain application. The major contribution of this paper is to apply SVMs technique to separate VGCNF/VE nanocomposite test data into various desired mechanical property classes. As a result, an unknown VGNCF/VE specimen (*i.e*., a configuration not represented by the current dataset) can be easily characterized and classified into its corresponding VGCNF/VE class without the need to conduct expensive and time-consuming experiments. This quick qualitative assessment significantly reduces the lead time on developing a new material system for a desired application.

2 Materials and Methods

All data used in this work were generated using various statistical experimental designs, such as a general mixed level full factorial and central composite design, and are described in detail elsewhere [18-21, 27]. Different datasets were merged into a larger one incorporating 240 viscoelastic data points, 60 flexural data points, 172 compression data points, and 93 tension data points for variously formulated and processed VGCNF/VE nanocomposites. Therefore, the new larger dataset has a total of 565 data points. Each data point corresponds to combinations of nine input design factors and nine output responses. The input factors of the new VGCNF/VE dataset are curing environment (air vs. nitrogen), use or absence of a dispersing agent, strain rate, mixing method (ultrasonication, high-shear mixing, combination of both), VGCNF weight fraction, VGCNF and type (pristine vs. oxidized), high-shear mixing time, sonication time, and temperature. The output factors (*i.e*., measured properties) are true ultimate strength, true yield strength, engineering elastic modulus, engineering ultimate strength, flexural modulus, flexural strength, storage modulus, loss modulus, and tan delta. Therefore, the effectiveness of the SVMs technique implemented in this study is that materials scientists and engineers can select the optimal manufacturing combination of input factors that yield a desired mechanical property response; namely high storage modulus response, high true ultimate strength response, or high flexural modulus response. The choice of the optimal combination is based on several industrial measures, among which is the inputs' combination that has the minimum fabrication cost, the fastest or the most time-efficient combination, the combination that results in the best mechanical properties of the resulting VGCNF/VE nanocomposites, or a combination of two or more of these measures.

Different data interpolation techniques were used to replace some of the missing and unknown data fields in the new dataset [28]. These techniques include linear interpolation which is a method of curve-fitting using linear polynomials, and spline interpolation where the interpolant is a spline (piecewise polynomial). However, spline interpolation is more precise than regular polynomial interpolations because of its low interpolation error regardless of the polynomial degree used for the spline. In addition, spline interpolation avoids the problem of Runge's phenomenon, which occurs when using high degree polynomials for the interpolation process [28].

3 Theory/Calculation

As mentioned in Section 2, this study incorporates nine input and nine output design factors. Therefore, the dataset represents an eighteen-dimensional (18-D) analysis case. Since curing environment, use or absence of dispersing agent, mixing method, and VGCNF type are considered qualitative factors, they are represented by a numeric code for the analysis purposes. All quantitative values were normalized using standardized scores since the original value ranges were dissimilar.

Resubstitution and 3-fold repetitive cross validation techniques were used with the dataset to characterize the specimens that have desired VGCNF/VE properties. Each specimen was separated into an appropriate VGCNF/VE mechanical property class: specimens with high storage modulus (class 1), specimens with high true ultimate strength (class 2), and specimens with high flexural modulus (class 3). Before applying these techniques, a brief explanation of the SVMs operations, resubstitution, and repetitive cross validation techniques are introduced.

3.1 SVMs Operations

The goal of an SVMs classifier is to find a separating hyperplane between the points belonging to two distinct classes and maximize the distance between these points to the hyperplane. This maximum distance is referred to as the margin. This concept is illustrated in Figure 1 [1] for linearly separable data. For nonlinearly separable data, the resulting hyperplane and margin has a complex, nonlinear form as shown in Figure 2 [1].

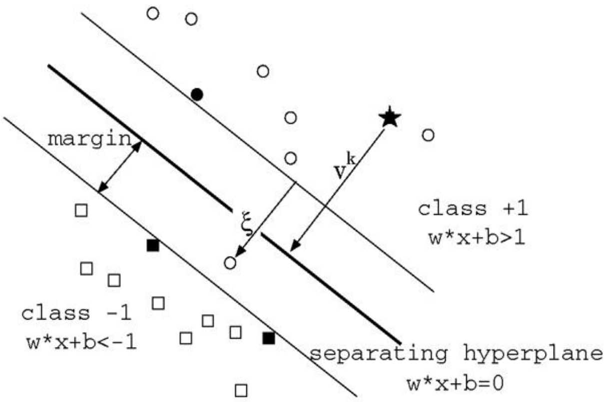

Figure 1. The SVMs model: the separating hyperplane along with the maximum margin for linearly separable data [1].

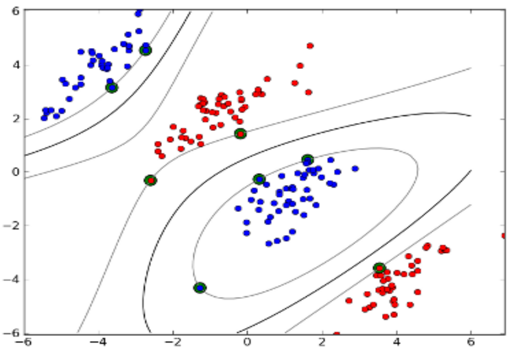

Figure 2. An example of the SVMs model for nonlinearly separable data [1].

3.2 Resubstitution method

The resubstitution method [29] is a computationally efficient technique in which the whole dataset is used to train the SVMs model, and the same dataset is used for testing (validation). This ensures that the SVMs model generalizes well when combinations of inputs and outputs are applied whose classes are not explicitly known. Good generalization is achieved when the apparent error (AE) is minimized [24].

The AE is defined as:

$$AE = \frac{1}{N}\sum_{i=1}^{N}\left|t_i - a_i\right| \quad (1)$$

where N is the total number of samples, t_i is the targeted class of the sample in binary classification (i.e. 1 if the sample belongs to one class and 0 if it belongs to the other class), and a_i is the actual SVMs binary classification value (0 or 1).

Although several SVMs architectures and training algorithms are available, the SVMs classifier for two nonlinearly separable data is the most commonly used one and was utilized in this study [1]. However, since this study deals with separating the VGCNF/VE specimens into three different distinct property classes, the designed SVMs model was implemented in three stages using a one-against-all (OAA) strategy [30]. For example, in the first stage, specimens belonging to class 2 and class 3 were combined and compared against class 1. Finally, the classification information from these stages was combined in order to determine the three distinct property classes. This SVMs model assumed a non-linear relationship between the input-output variables and the corresponding class associated with each sample.

3.3 Repetitive cross validation technique

Repetitive cross validation (RCV) techniques can better train the SVMs model using available data. First, the available dataset is randomly partitioned into a training set and a test set. The training samples are further partitioned into two disjoint subsets: 1) the estimation subset, which is used to select the SVMs, and 2) the validation subset, which is used to test or validate the developed SVMs model [31]. In this way, the training samples can be used to assess the performance of various candidate SVMs models and thus the “best” one can be chosen [31]. Currently, there are four different RCV methods: holdout RCV, early-stopping method of training, multi-fold RCV, and leave-one-out RCV [32].

4 Results and discussion

The workflow of the classification process begins with applying the developed SVMs model to the VGCNF/VE data in which it was divided into training and test sets. Then two techniques were used for performance evaluation of the SVMs classifier; the resubstitution and the 3-folds RCV techniques. Finally, the results from both techniques were compared. In essence, the classifier ability to identify the percentage of test samples that belong to each of the desired mechanical properties (high storage modulus, high true ultimate strength, and high flexural modulus) was evaluated and analyzed.

In the SVMs analyses, classifications were compared and analyzed by using sets of confusion matrices (contingency tables) [33]. Additionally, a positive constant, *C*, was used to balance the margin size and the misclassification instances. The choice for *C* determines the number of support vectors and the overall performance of the SVMs model [1]. Three values of C were used in this study: 0.5, 10, and 100. Three kernel functions were also used in this study: a degree two polynomial, a dot product, and a hyperbolic tangent kernels.

For the resubstitution method, the SVMs model was generally able to correctly classify 100% of the VGCNF/VE specimens into three different distinct property classes for all kernel functions used in this study regardless of the constant *C*. Although a minor classification error (5%) resulted when the hyperbolic tangent kernel was used for class 3, this error was considered to be acceptable as it did not affect the overall classification accuracy of the model. These high classification rates are due to the fact that all samples were used for training and testing so the amount of misclassified information was minimal. For the 3-fold RCV technique, the chosen sizes of training and testing sets were 80 and 40 data samples, respectively, for each of the three folds.

Following the standard practice of the SVMs analysis, the inputs and outputs were normalized using standardized scores, as their original value ranges were completely different from each other. The classification performance of the 3-folds RCV technique was inferior to that of the resubstitution method. Since the class 1, class 2, and class 3 sizes for the three folds in all stages were significantly lower than such classes when the resubstitution method was applied, some misclassification error is to be expected. Also, unlike the resubstitution method, the same samples were not used for training and testing in each fold resulting in some additional misclassification error. However, the resulted confusion matrices

Table 1. Classification information of the SVMs model when a polynomial kernel of degree 2 was implemented using both the resubstitution and 3-fold RCV methods.

	Resubstitution method				**3-fold RCV method**			
	Class 1	Class 2	Class 3	Average	Class 1	Class 2	Class 3	Average
Correct Classification Rate	100%	100%	100%	100%	100%	75.00%	33.33%	69.44%
Apparent Error Rate	0%	0%	0%	0%	0%	25.00%	66.67%	30.56%

Table 2. Classification information of the SVMs model when a dot product kernel was implemented using both the resubstitution and 3-fold RCV methods.

	Resubstitution method				**3-fold RCV method**			
	Class 1	Class 2	Class 3	Average	Class 1	Class 2	Class 3	Average
Correct Classification Rate	100%	100%	100%	100%	100%	50.00%	47.50%	65.83%
Apparent Error Rate	0%	0%	0%	0%	0%	50.00%	52.50%	34.17%

Table 3. Classification information of the SVMs model when a hyperbolic tangent kernel was implemented using both the resubstitution and 3-folds RCV methods.

	Resubstitution method				**3-fold RCV method**			
	Class 1	Class 2	Class 3	Average	Class 1	Class 2	Class 3	Average
Correct Classification Rate	100%	100%	95.00%	98.33%	98.75%	50.00%	81.67%	76.81%
Apparent Error Rate	0%	0%	5.00%	1.67%	1.25%	50.00%	18.33%	23.19%

showed that the SVMs classifier performed well for fold 3 samples for all kernel functions at about 100% classification rate. In addition, reasonable classification rate was achieved for fold 2 when the hyperbolic tangent kernel was implemented at 75.00% and 58.33% was obtained for fold 2 samples when the polynomial kernel (degree 2) was implemented. The classification rates were lower for other cases. In addition, similar to the resubstitution method analyses, the classification results were independent of the value of the constant *C*. Another observation is that while 3-folds RCV technique was able to correctly classify specimens into class 1, mixed results were obtained when classifying specimens according to class 2 and class 3 and were observed to be dependent on the kernel function. For example, when a hyperbolic tangent kernel, the correct classification rate for class 3 was observed to be 81.67%. This value dropped to 33.33% when the degree two polynomial kernel was used. On average, the classification performance was the best when hyperbolic tangent kernel was implemented yielding a classification rate of 76.81%.

The overall classification rates and apparent error rates for the three different kernels using both the resubstitution and 3-fold RCV methods are shown in Tables 1-3. Based on these results, the SVMs model was able to more correctly classify specimens belonging to class 1 than class 2 and 3. Additionally, the resubstitution method was determined to be superior to the 3-fold RCV method for this specific problem. For example, when the resubstitution method was implemented using the hyperbolic tangent kernel, the SVMs model was able to identify all samples (100%) that have the highest storage modulus responses and all samples that have the highest true ultimate strength responses where it was able to identify 95% of samples that have the highest flexural modulus responses (Table 3). When the 3-fold RCV was implemented, the SVMs model was able to identify 98.75% of test samples that have the highest storage modulus responses, 50% of test samples that have the highest true ultimate strength responses, and 81.67% of test samples that have the highest flexural modulus responses (Table 3).

In addition, by choosing particular inputs' combination based on one of the industrial optimal measures mentioned in section 2, this SVMs model is able to identify the desired mechanical property (one of the three desired mechanical response classes of high storage modulus, high true ultimate strength, or high flexural modulus) that will be resulted from this combination based on the selected industrial measure. Section 5 will elaborate more on the effectiveness of the developed SVMs model on the VGCNF industrial manufacturing process.

5 Application to Materials Informatics

Since this dataset encompasses a wide variety of mechanical testing methods, conditions, and material configurations, the resulting SVMs model can be used to effectively predict the mechanical response for previously untested material configurations. Such a capability reduces the need to perform further experiments and allows the materials scientists to quickly assess the viability of a new material configuration. For example, if a high storage modulus is desired, the optimal VGCNF weight fraction can be determined for given mixing conditions which are likely not located at one of the tested levels. Additionally, material and processing costs can likely be reduced by using material informatics principles. Since VGCNFs are often expensive and nanocomposite fabrication processes often take several hours, a range of VGCNF weight fractions and mixing times can be established over which adequate properties are obtained. A smaller amount of VGCNFs combined with shorter mixing times could ultimately reduce production costs by a significant amount.

6 Summary and Conclusions

A support vector machines (SVMs) technique was applied to a vapor-grown carbon nanofiber (VGCNF)/vinyl ester (VE) nanocomposite dataset as a proof of concept for materials informatics. This dataset consists of 565 different design points: 172 compression, 93 tension, 60 flexure, and 240 viscoelastic points. Each treatment combination consisted of eighteen feature dimensions corresponding to the nine input and nine output design factors. The nine input factors of the VGCNF/VE dataset were curing environment (air vs. nitrogen), use or absence of a dispersing agent, strain rate, mixing method (ultrasonication, high-shear mixing, and combination of both), VGCNF weight fraction, VGCNF type (pristine vs. oxidized), high-shear mixing time, sonication time, and temperature. The output factors (*i.e.*, measured properties) were true ultimate strength, true yield strength, engineering elastic modulus, engineering ultimate strength, flexural modulus, flexural strength, storage modulus, loss modulus, and tan delta. The SVMs model was trained using the resubstitution method and the 3-fold repetitive cross validation (RCV) technique to classify each VGCNF/VE sample into one of three optimal property classes: high storage modulus, high true ultimate strength, high flexural modulus. The classifier was implemented in three stages using a one-against-all strategy. Three possible kernel functions were explored in this study: a polynomial kernel of degree two, a dot product kernel, and a hyperbolic tangent kernel. A set of confusion matrices was used to compare the different analysis methods.

In general, the SVMs model using the resubstitution method was able to predict the optimal property classes with a minimal apparent error (AE) rate irrespective of the kernel function that was used. While the SVMs model using the 3-fold RCV method was able to accurately predict which data points belonged to the high storage modulus class, in general, the SVMs model using this method had significant FARs.

Most importantly, the developed SVMs model is able to identify the desired mechanical property response value (high storage modulus, high true ultimate strength, or high flexural modulus) resulted from a chosen untested combination of the nine input factors mentioned in this study. The choice of the inputs' combinations commensurate with particular optimal industrial measure(s) selected by materials scientists and engineers. This includes but is not limited to, the inputs' combination that has the minimum industrial fabrication cost, the combination that yields the fastest manufacturing process, the combination that results in the optimal mechanical properties of the resulting VGCNF/VE polymer nanocomposites, or any two or more of these measures combined. In other words, if an inputs combination whose outputs responses are unknown is given to the developed SVMs model, then the desired mechanical property response will be easily retrieved based on this combination.

The model's ability to identify these desired mechanical property responses based on particular combination of input factors will result in faster VGCNF nanocomposites manufacturing lead time without the need to rely on intensive and time-consuming experiments. In addition, while this model only considers three classes, it can also be readily extended to include additional desirable material properties. This issue is the focus of ongoing research.

The SVMs classifier applied in this study demonstrates the usefulness of data mining and knowledge discovery techniques in materials science and engineering. It is expected that more such techniques will be employed within the rising field of materials informatics in near future.

References

[1] S. Theodoridis, K. Koutroumbas, Pattern Recognition, fourth ed., Academic Press, Massachusetts, 2008.

[2] K. Rajan "Materials informatics," *Materials Today*. Vol. 8, pp. 38-45, 2005.

[3] K. F. Ferris, L. M. Peurrung, J. M. Marder "Materials informatics: Fast track to new materials," *Advanced Materials and Processes*, Vol. 165, pp. 50-51, 2007.

[4] C. Suh, K. Rajan, B. M. Vogel, B. Narasimhan, S. K. Mallapragada "Informatics methods for combinatorial materials science," *Combinatorial Materials Science*, Vol. 5, pp. 109-119, 2006.

[5] Q. Song "A preliminary investigation on materials informatics," *Chinese Science Bulletin*, Vol. 49, pp. 210-214, 2004.

[6] J. Rodgers and D. Cebon (2006). Materials Informatics. *MRS Bulletin*, 31, pp 975-980. doi:10.1557/mrs2006.223.

[7] AbuOmar, O., Nouranian, S., & King, R. (2013). Artificial Neural Network Modeling of the Viscoelastic Properties of Vapor-Grown Carbon Nanofiber/Vinyl Ester Nanocomposites. *The 19th International Conference on Composite Materials (ICCM19)*, July 28-August 2, 2013. Montreal, Quebec, Canada.

[8] Roberts, K, Muchlich, F, Schenkel, R, Weikum, G, An Information System for Material Microstructures. *International Conference on Scientific and Statistical Database Management (SSDBM'04)*. 1099-3371/04, 2004 IEEE

[9] Swaddiwudhipong, S, Tho, K K, Liu, Z S, Hua, J, Ooi, N S B. Material Characterization via Least Squares Support Vector Machines. *Modelling Simul. Mater. Sci. Eng*. 13 (2005) 993-1004.

[10] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. DeMoor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, 2002 (ISBN 981-238-151-1)

[11] C. Hu, C. Ouyang, J. Wu, X. Zhang, C. Zhao "NON-structured materials science data sharing based on semantic annotation," *Data Science*. Vol. 8, pp. 52-61, 2009.

[12] T. Sabin, C. Bailer-Jones, P. Withers "Accelerated learning using Gaussian process models to predict static recrystallization in an Al-Mg alloy," *Modelling and Simulation in Materials Science and Engineering*, Vol. 8, pp. 687-706, 2000.

[13] J. H. Koo "*Polymer nanocomposites: Processing, characterization, and applications*," 1st edition, McGraw-Hill, New York, 2006.

[14] J. Garces, D. J. Moll, J. Bicerano, R. Fibiger, D. G. McLeod "Polymeric nanocomposites for automotive applications," *Advanced Materials*, Vol. 12, pp. 1835-1839, 2000.

[15] F. Hussain, M. Hojjati, M. Okamoto, R.E. Gorga, Review article: polymer-matrix nanocomposites, processing, manufacturing, and application: An overview, J. Compos. Mater., 40 (2006) 1511-1575.

[16] E.T. Thostenson, C. Li, T.W. Chou, Nanocomposites in context, Compos. Sci. Technol., 65 (2005) 491-516.

[17] O. Abuomar, S. Nouranian, R. King, J.L. Bouvard, H. Toghiani, T. E. Lacy, C. U. Pittman Jr "Data mining and knowledge discovery in materials science and engineering: A polymer nanocomposites case study," *Advanced Engineering Informatics* 27(4), 615–624, 2013. DOI:10.1016/j.aei.2013.08.002.

[18] S. Nouranian, Vapor-grown carbon nanofiber/vinyl ester nanocomposites: Designed experimental study of mechanical properties and molecular dynamics simulations, Mississippi State University, PhD Dissertation, Mississippi State, MS USA, 2011.

[19] S. Nouranian, H. Toghiani, T.E. Lacy, C.U. Pittman, J. Dubien, Dynamic mechanical analysis and optimization of vapor-grown carbon nanofiber/vinyl ester nanocomposites using design of experiments, J. Compos. Mater., 45 (2011) 1647-1657.

[20] S. Nouranian, T.E. Lacy, H. Toghiani, C.U. Pittman Jr, J.L. Dubien, Response Surface Predictions of the Viscoelastic Properties of Vapor-Grown Carbon Nanofiber/Vinyl Ester Nanocomposites, J. Appl. Polym. Sci., (2013) DOI: 10.1002/app.39041.

[21] G.W. Torres, S. Nouranian, T.E. Lacy, H. Toghiani, C.U. Pittman Jr, J. Dubien, Statistical Characterization of the Impact Strengths of Vapor-Grown Carbon Nanofiber/Vinyl Ester Nanocomposites Using a Central Composite Design, J. Appl. Polym. Sci., 128 (2013) 1070-1080.

[22] R.L. King, A. Rosenberger, L. Kanda, Artificial neural networks and three-dimensional digital morphology: a pilot study, Folia Primatol., 76 (2005) 303-324.

[23] T. Kohonen, Self-organization and associative memory, Springer-Verlag, 1988.

[24] S. Miyamoto, H. Ichihashi, K. Honda, Algorithms for Fuzzy Clustering: Methods in c-Means Clustering with Applications, Springer, 2008.

[25] J.C. Bezdek, R. Ehrlich, FCM: The fuzzy c-means clustering algorithm, Computers & Geosciences, 10 (1984) 191-203.

[26] I.T. Jolliffe, Principal Component Analysis, Springer, 2002.

[27] J. Lee, S. Nouranian, G.W. Torres, T. E. Lacy, H. Toghiani, C. U. Pittman, J. L. DuBien, Characterization, prediction, and optimization of flexural properties of vapor-grown carbon nanofiber/vinyl ester nanocomposites by response surface modeling. J. Appl. Polym. Sci., 130 (2013) 2087–2099. doi: 10.1002/app.39380

[28] MATLAB Mathematics and Interpolation, Release 2012a, The MathWorks, Inc., Natick, Massachusetts, United States.

[29] J. Twomey, A. Smith "Bias and variance of validation methods for function approximation neural networks under conditions of sparse data," *IEEE Transactions on Systems, Man., and Cybernetics*, Vol. 28, pp. 417-430, 1998.

[30] J. Milgram, M. Cheriet, and R. Sabourin, "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?." *Tenth International Workshop on Frontiers in Handwriting Recognition*. 2006.

[31] S. Haykin, Neural networks and learning machines, 3rd edition, Prentice-Hall, 2009.

[32] S. Hayken, Neural networks: A comprehensive foundation, 2nd edition, Englewood Cliffs (NJ): Prentice-Hall, 1999.

[33] R. Kovari, F. Provost, Glossary of terms, *Machine Learning*, 30 (1998), No. 2/3, 271-274.