

## Designing Structural Parameters of Nonwovens Using Fuzzy Logic and Neural Networks

PHILIPPE VROMAN, LUDOVIC KOEHL, XIANYI ZENG

*Ecole Nationale Supérieure des Arts et Industries Textiles, 9 rue de l'Ermitage  
59100 Roubaix, France*

TING CHEN

*College of Textile and Clothing Engineering,  
Soochow University, Suzhou 215021, China*

Received: 18-06-2008

Revised: 30-09-2008

### Abstract

In this paper, a computer aided system for designing nonwoven materials is presented. As an original approach in the field of nonwoven research, both quality measurement analysis and human knowledge processing are integrated in the system. It allows designers to optimize the structure of nonwoven materials with limited trials according to the functional properties given in customers' specifications. This system aims at modeling the relation between functional or physical properties (outputs) and structural parameters (inputs) of nonwoven products. In order to reduce the complexity of the system, a procedure is proposed for selecting the most relevant input variables based on a ranking criterion, which takes into account both the expertise of manufacturers and the measured data. In this criterion, fuzzy logic is used to establish a good compromise or a fusion between these two uncertain and incomplete information sources. Then, two models are set up by utilizing multilayer feed forward neural networks, which take into account the generality and the specificity of the product families respectively. The presented models have been validated with the use of experimental data concerning several families of nonwoven products.

*Keywords:* nonwoven material structures, physical properties, fuzzy-neural design support system, parameter selection

### 1. Introduction

Nonwoven products are fibrous materials characterized by a large range of interesting properties, due mainly to the diversity of raw materials, forming, bonding and finishing technologies. As durable or semi-durable materials, they are employed in various application fields such as manufacturing, civil engineering, building or transportation. Their increasing success is also due to the good ratio performance/cost price. Consequently, the number of end-products designed with nonwoven materials has significantly grown in the last decades while the production in Western Europe has risen by 8%.<sup>1</sup>

Due to the international competition in the textile market, nonwoven materials should be designed and produced in order to satisfy more and more complex

specifications (e.g. insulation, protection, filtration, durability, breathiness...) and increasing requirements for international standards in different application fields. In parallel, nonwoven product designers are actively involved in projects to reduce cost by applying value analysis during the design and the development of these manufacturing products. Several criteria of the product design are given as follows.

- 1) Satisfying all specific values of the functional properties of nonwoven materials (i.e. customer's specifications),
- 2) Minimizing the cost and the quantity of raw materials,
- 3) Optimizing the final structure of materials.

Recently, great attention has been paid to explore the relationship between the structural parameters of nonwoven materials (thickness, basis weight, raw

Philippe Vroman, Ludovic Koehl, Xianyi Zeng, Ting Chen

material, number of layers...) and their functional properties. This approach enables nonwoven manufacturers to obtain a better understanding of the effect of the material structure and the corresponding process parameters on the product quality.

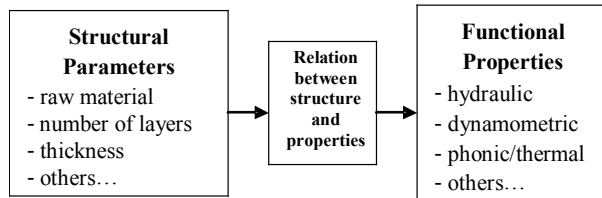


Fig. 1. Modeling the relations between structural parameters and functional properties of nonwovens.

The aim of our project is to develop a design support system for product designers using fuzzy logic and neural networks. The system includes a number of mathematical models for characterizing the relations between the structural parameters (input variables) and the functional properties (output variables) in order to optimize material structure and predict quality of new nonwoven products<sup>2</sup> (Fig. 1). However, this procedure of modeling is very complex because of the nonlinear relationship between inputs and output variables, the large number of structural parameters, the interdependencies between them and the critical lack of available learning data. In practice, the amount of learning data or learning samples is strongly constrained by the production costs or experiment costs. Moreover, the production lines are not always available for trials. Given these constraints, a small set of learning samples have been used to model the relationship between structural parameters and functional properties of materials.

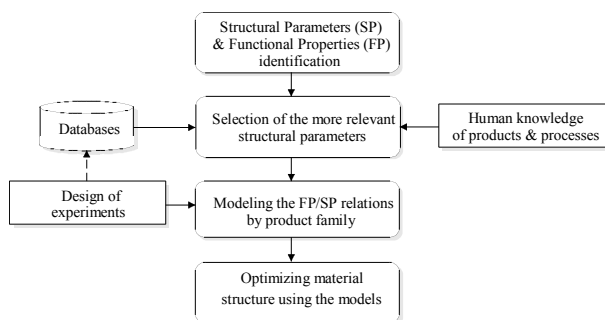


Fig. 2. Framework of the proposed design support system.

The structure of the proposed support system for designing new nonwoven products is shown in Fig.2. In this system, an exhaustive list of structural parameters (SP) and functional properties (FP) related to a specific application problem are firstly extracted according to the physical knowledge of nonwoven experts. Then, the most relevant structural parameters are selected from the list by combining the measured data obtained from a small number of experiments and the human knowledge of operators on processes and products. Fuzzy logic is used in the generation of this selection criterion in order to find a good compromise or a fusion between these two information sources. Finally, for each product family, a neural network is set up for modeling the relationship between the selected structural parameters and the concerned functional property. Based on this model, designers can optimize the structure of the nonwoven product according to the specifications.

This paper is organized as follows. In Section 2, a procedure is given for selecting the most relevant input variables (structural parameters) based on a ranking criterion for reducing the complexity of the models. This fuzzy logic based selection criterion has been developed by properly integrating both human knowledge of operators on processes and products and measured data. In Section 3, we present the modeling procedure for characterizing the relationship between the selected relevant structure parameters and each functional property of the nonwoven. Multilayer feed forward Artificial Neural Network (ANN) models<sup>3</sup> have been built with specific architectures adapted to the product diversity. In Section 4, the proposed models have been successfully applied to the prediction of hydraulic properties of filtration media nonwoven products. Finally, a general conclusion is given in Section 5.

## 2. Selection of Relevant Structural Parameters

### 2.1. Analysis of the existing work

When studying the effect of each structural parameter on the functional properties selected from the final product specifications, it is quite difficult to produce a large number of samples. Therefore, small-scaled ANN models are built from a limited number of learning data and the most relevant structural parameters are selected before the modeling procedure.

The existing algorithms of variable selection mostly use heuristic search for an optimal subset of the original variables, with each state in the search space specifying a subset of the possible variables. Generally, each variable selection method is designed according to the following basic issues that determine the nature of the heuristic process<sup>4</sup>:

- 1) The direction of search and the operators used. Two methods are frequently used: forward selection and backward elimination.
- 2) The organization of the search.
- 3) The strategy used to evaluate alternative subsets of variables.
- 4) The condition for halting the search.

In literature, most of the variable selection methods deal with data based classification problems. Moreover, the strategy used for evaluating variables and the condition for halting the search are generally defined as the variable's ability to discriminate among classes of the learning data. The optimal subset of variables corresponds to the case in which the separability between different classes is maximal and data inside each class are as compact as possible.

Most of the existing work has been carried out in the frame of supervised variable selection, i.e. the objective of selection is to improve the classification accuracy or class label predictive accuracy of data samples.<sup>5</sup> Several well-known methods are the decision-tree method,<sup>6</sup> the nearest-neighbor method,<sup>7</sup> the mutual information measure based method<sup>8</sup> and the hyperbox generation based method,<sup>9</sup> the information-theoretical connectionist network model for removing both irrelevant and redundant variables<sup>10</sup> and the wrapper model, which evaluates alternative subsets of variables by running some induction algorithm on the learning data and using the estimated accuracy of the resulting classifier as its metric.<sup>11</sup> There also exists some work on unsupervised variable selection using conditional Gaussian networks<sup>5</sup> and data clustering techniques.<sup>12</sup> Recently, a new variable selection method has been developed using a modified fuzzy C-means algorithm with supervision.<sup>13</sup>

In practice, the performance of these data based variable selection methods is strongly related to the quality and the quantity of data samples and the criterion defined,

which may vary from task to task. These methods are not efficient to solve variable selection problems in some industrial processes. In these processes, limited by the cost and the time of measurement, the quantity of data is often too small to constitute a correct distribution for obtaining significant classification results. In this case, the class separability based criteria of variable selection should be replaced by variable sensitivity based criteria such as gradient descent. Moreover, if possible, physical knowledge related to the problem and measured numerical data should be used in a complementary way in order to improve the criterion of selection and cross-validate the results obtained from these two information sources.

## 2.2. Formalization of the criterion for relevant variable selection

In this paper, we first propose a criterion for ranking the nonwoven structural parameters by linearly combining the human knowledge based criterion and the data sensitivity to the properties. The related formalization is given below.

Let  $m$  and  $n$  be the total number of structural parameters and the total number of functional properties respectively. The input and output variables are denoted as  $X = \{x_1, x_2, \dots, x_m\}$  and  $\{y_1, y_2, \dots, y_n\}$  respectively. The relationship between  $\{x_1, x_2, \dots, x_m\}$  and one output variable  $y_l$  can be considered as a nonlinear function  $f$  so that  $y_l = f(x_1, x_2, \dots, x_m)$ . For a  $t$  sized subset of input variables in  $X$  denoted as  $X^t = \{x_{(1)}, x_{(2)}, \dots, x_{(t)}\}$ , we create a new nonlinear function  $g(x_{(1)}, x_{(2)}, \dots, x_{(t)})$  in which we aggregate the variables of  $X - X^t$  by calculating the average of all values of  $f(x_1, x_2, \dots, x_m)$  for these remaining variables. The elements in the  $t$ -sized subset  $\{x_{(1)}, x_{(2)}, \dots, x_{(t)}\}$  are considered as the  $t$  most relevant variables if and only if their mean value of  $|f - g|$  is the smallest for all the  $t$  sized subsets of input variables in  $X$ .

Let  $X_s = (x_{s1}, x_{s2}, \dots, x_{sb}, \dots, x_{sm})^T$  and  $Y_s = (y_{s1}, y_{s2}, \dots, y_{sb}, \dots, y_{sn})^T$  be the input vector of structural parameters and the output vector of functional properties that correspond to the sample  $s$  ( $s \in \{1, \dots, z\}$ ) respectively. All the measured data have been normalized to eliminate the scale effects and the learning data set contains  $z$  samples. In order to rank the relevant inputs for a given output  $y_l$ , a criterion variable  $F_{k,l}$  (estimated

Philippe Vroman, Ludovic Koehl, Xianyi Zeng, Ting Chen

for each input variable  $x_k$  and related to output variable  $y_l$ ) is defined as follows:

$$F_{k,l} = g_1 \cdot H_{k,l} + g_2 \cdot S_{k,l} \quad (1)$$

with  $k \in \{1, \dots, m\}$ ,  $l \in \{1, \dots, n\}$ ,  $g_1$  and  $g_2$  are two positive coefficients.

The criterion  $F_{k,l}$  is designed for searching the best compromise between the sensitivity variation of measured data and the conformity of the human knowledge to measured data, represented by  $S_{k,l}$  and  $H_{k,l}$  respectively. The larger  $F_{k,l}$  is, the more relevant the input  $x_k$  is to the output  $y_l$ .

### 2.3. Sensitivity of measured data

The sensitivity is a distance based criterion for evaluating the effects of the input variables on the output variable. It is defined according to the following two assumptions:

- 1) IF a small variation of the input variables corresponds to a large variation of the output variable, THEN the sensitivity of these variables is important.
- 2) IF a large variation of an input variable corresponds to a small variation of the output variable, THEN the sensitivity of these variables is not important.

The criterion of sensitivity for all input variables related to the output variable  $y_l$  is defined by

$$S_l = \frac{2}{\pi} \arctan \left( \sum_{\substack{i,j \in \{1, \dots, z\} \\ i \neq j}} \frac{d(y_{il}, y_{jl})}{d(X_i, X_j)} \right) \quad (2)$$

where  $d(X_i, X_j)$  is the Euclidean distance between two input data  $X_i$  and  $X_j$  and  $d(y_{il}, y_{jl})$  the Euclidean distance between  $y_{il}$  and  $y_{jl}$ .

Evidently,  $S_l$  varies between 0 and 1. This criterion is used for evaluating the sensitivity of the whole set of input variables related to the output variable  $y_l$ . If values of  $S_l$  are close to 1, then we consider that small variations of input data can cause big variations in the output space and the input variables are sensitive to measured data. If the values of  $S_l$  are close to 0, then we consider that big variations of input data correspond to small variations in the output space and the input variables are insensitive to measured data.

The criterion  $S_l$  can be considered as a measure of information content in the input variables. However, for selecting relevant variables, we need to evaluate the information content after removing one or a group of input variables. The criterion of sensitivity related to the output variable  $y_l$  when removing the input variable  $x_k$  is then defined by

$$S_{k,l} = \frac{2}{\pi} \arctan \left( \sum_{\substack{i,j \in \{1, \dots, z\} \\ i \neq j}} \frac{d(y_{il}, y_{jl})}{d'_k(X_i, X_j)} \right) \quad (3)$$

where  $d'_k(X_i, X_j) = \sqrt{d^2(X_i, X_j) - d_k^2(X_i, X_j)}$  and  $d_k(X_i, X_j)$  is the projection of  $d(X_i, X_j)$  on the axis  $x_k$ .

From Eq. (2) and Eq. (3), we can easily obtain  $1 > S_{k,l} > S_l$ . For a specific input variable  $x_k$ , if the value of  $S_{k,l}$  is bigger than any other sensitivity value  $S_{p,l}$  after removing related input variable  $x_p$  ( $p \neq k$ ), then we consider that  $x_k$  is the most insensitive to measured data because the remaining input variables after removing  $x_k$  are more sensitive than those after removing any other individual input variable. According to the same idea, if the value of  $S_{k,l}$  is smaller related to the other input variables, then we consider that  $x_k$  is the most sensitive to measured data because the remaining input variables after removing  $x_k$  are the least sensitive.

In order to be conform to the definition of  $F_{k,l}$  in Eq. (1), i.e. big values of sensitivities correspond to relevant variables and small values of sensitivities to irrelevant variables, we transform Eq. (3) into the following form:

$$S_{k,l} = 1 - \frac{2}{\pi} \arctan \left( \sum_{\substack{i,j \in \{1, \dots, z\} \\ i \neq j}} \frac{d(y_{il}, y_{jl})}{d'_k(X_i, X_j)} \right) \quad (4)$$

Evidently, values of  $S_{k,l}$  also vary between 0 and 1. The bigger the value of the sensitivity criterion  $S_{k,l}$  is, the more sensitive the corresponding variable  $x_k$  is to measured data.

### 2.4. Conformity of human knowledge to measured data

In Eq. (1),  $H_{k,l}$  represents the degree of coherence between the human knowledge expressed in Table 1 and

the variation of measured data. Its principle is given as follows. If a variable  $x_k$  has the same variation trend in learning data set as in the human knowledge, it will be considered as relevant. Otherwise, it will be considered as irrelevant. The universe of discourse of  $y_l$  is divided into  $t$  equal intervals  $C_{lp}$  ( $p=1, \dots, t$ ). The set  $A_{kp}$  is constructed using the projection of the input data set on the axis  $x_k$ , which corresponds to the output interval  $C_{lp}$  of  $y_l$ .  $I_{kp}$  is generated by the overlap between  $A_{kp}$  and  $A_{kp+1}$  ( $p \in \{1, \dots, t-1\}$ ) (see Fig. 3).

Table 1. Formalization of the human knowledge table.

Structural parameters (Input space)	End-use functional properties (Output space)				
	$y_1$	...	$y_t$	...	$y_n$
$x_1$	$R(x_1, y_1)$	...	$R(x_1, y_t)$	...	$R(x_1, y_n)$
...	...	...	...	...	...
$x_k$	$R(x_k, y_1)$	...	$R(x_k, y_t)$	...	$R(x_k, y_n)$
...	...	...	...	...	...
$x_m$	$R(x_m, y_1)$	...	$R(x_m, y_t)$	...	$R(x_m, y_n)$

positive influence:  $R(x_i, y_j) = +1$

negative influence:  $R(x_i, y_j) = -1$

no influence:  $R(x_i, y_j) = 0$

no human knowledge: Empty cell

The human knowledge is summarized in Table 1. It is provided by a number of operators working on the related processes of nonwoven materials according to their experience. In practice, they have some incomplete qualitative knowledge on variation trends of end-use functional properties with non woven structural parameters. Each cell of Table 1 corresponds to one IF...THEN rule relating the input  $x_k$  to the output  $y_l$ . Several examples are given below.

IF  $x_k$  is increasing THEN  $y_l$  is increasing:  $R(x_k, y_l) = +1$

IF  $x_k$  is increasing THEN  $y_l$  is decreasing:  $R(x_k, y_l) = -1$

There is no influence of  $x_k$  on  $y_l$ :  $R(x_k, y_l) = 0$

The element  $H_{k,l}$  can then be calculated using the following formula:

$$\begin{cases}
 H_{k,l} = \frac{1}{t-1} \sum_{p=1}^{t-1} h_p \\
 x_{kp}^{\inf} = \min_{s \in \{1, \dots, z\}} \{x_{sk} | y_{sl} \in C_{lp}\} \text{ and } x_{kp}^{\sup} = \max_{s \in \{1, \dots, z\}} \{x_{sk} | y_{sl} \in C_{lp}\} \\
 \text{if } I_{kp} = \Phi, \begin{cases} h_p = \frac{1}{2} |R(x_k, y_l)| \times [1 + R(x_k, y_l)] & \text{if } x_{kp+1}^{\inf} \geq x_{kp}^{\sup} \\ h_p = \frac{1}{2} |R(x_k, y_l)| \times [1 - R(x_k, y_l)] & \text{if } x_{kp+1}^{\sup} \leq x_{kp}^{\inf} \end{cases} \\
 \text{if } I_{kp} \neq \Phi, \begin{cases} h_p = \frac{1}{2} |R(x_k, y_l)| \times [1 + R(x_k, y_l)] \times \left(1 - \frac{|I_{kp}|}{|U_{kp}|}\right) & \text{if } x_{kp+1}^{\sup} \geq x_{kp}^{\sup} \\ h_p = \frac{1}{2} |R(x_k, y_l)| \times [1 - R(x_k, y_l)] \times \left(1 - \frac{|I_{kp}|}{|U_{kp}|}\right) & \text{if } x_{kp+1}^{\inf} \leq x_{kp}^{\inf} \end{cases}
 \end{cases}$$

(5)

As shown in Fig. 3,  $|I_{kp}|$  and  $|U_{kp}|$  are the lengths of the intervals, which respectively correspond to the intersection and union of  $A_{kp}$  and  $A_{kp+1}$  (input space), related to  $C_{lp}$  and  $C_{lp+1}$  (output space).

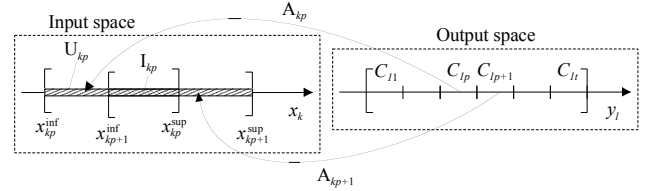


Fig. 3. Relationship between the input and output spaces.

Eq. (5) can be interpreted as follows.  $h_p$  represents the degree of coherence between the human knowledge and the variation of measured data in the two neighboring intervals  $C_{lp}$  and  $C_{lp+1}$ . Its value varies between 0 and 1. If  $I_{kp} = \Phi$  and  $x_{kp+1}^{\inf} \geq x_{kp}^{\sup}$ , then the input data  $x_s$ 's ( $s=1, \dots, z$ ) projected on the axis  $x_k$  are increasing when their corresponding output data  $y_{sl}$ 's vary from  $C_{lp}$  to  $C_{lp+1}$ . In this case, if the human knowledge  $R(x_k, y_l) = 1$  ( $x_k$  has a positive influence on  $y_l$ ), then we consider that this human knowledge is strongly coherent with the data variation trend of  $x_k$  in the two neighboring intervals  $C_{lp}$  and  $C_{lp+1}$ . From Eq. (5), we obtain  $h_p = 1$  (the best case). If the human knowledge  $R(x_k, y_l) = -1$  ( $x_k$  has a negative influence on  $y_l$ ), then we consider that this human knowledge is strongly incoherent with the data variation trend of  $x_k$  in the two neighboring intervals  $C_{lp}$  and  $C_{lp+1}$ . Then, we obtain  $h_p = 0$  from Eq. (5) (the worst case). Similar interpretation can be given to the case of  $I_{kp} = \Phi$  and  $x_{kp+1}^{\sup} \leq x_{kp}^{\inf}$ . If  $I_{kp} \neq \Phi$  and  $x_{kp+1}^{\sup} \geq x_{kp}^{\sup}$ , then we obtain a situation between the above two extreme cases (see Fig. 3) and the input data projected on  $x_k$  are slightly increasing when their corresponding output data vary from  $C_{lp}$  to  $C_{lp+1}$ . If  $R(x_k, y_l) = 1$ , then we consider that the human knowledge is weakly coherent with the data variation of  $x_k$  in  $C_{lp}$  and  $C_{lp+1}$ . The degree of coherence  $h_p$  is related to the overlap between the two data sets of  $x_k$  corresponding to  $C_{lp}$  and  $C_{lp+1}$ . The smaller this overlap is, the closer the data variation of  $x_k$  is to a strongly increasing case and the closer the value of  $h_p$  is to 1. If  $R(x_k, y_l) = -1$ , the human knowledge is incoherent with the data variation of  $x_k$  and we have  $h_p = 0$ . The other cases of  $h_p$  can be interpreted in the same way. If we obtain big values for all  $h_p$  ( $p=1, \dots, t-1$ ), then the value of the criterion  $H_{k,l}$  is also big and the

Philippe Vroman, Ludovic Koehl, Xianyi Zeng, Ting Chen

data variation related to  $x_k$  is coherent with the human knowledge.

After computing  $H_{k,l}$  and  $S_{k,l}$ , the value of the criterion  $F_{k,l}$  expressing the relevancy of each input  $x_k$  to a given output  $y_l$  can be determined. Then all the  $F_{k,l}$ 's ( $k=1, \dots, m$ ) can be ranked in a descending order. Consequently, the input corresponding to the highest value of  $F_{k,l}$  will be the most relevant input to this output, and so on.

### 2.5. Fuzzy logic based criterion for relevant variable selection

Two main drawbacks exist when using the raking criterion  $F_{k,l}$ .

- The coefficients  $g_1$  and  $g_2$  are difficult to be determined because the precise importance of the data sensitivity  $S_{k,l}$ , related to the coherence between human knowledge and measured data  $H_{k,l}$ , as well the precise ranges of these two variables, are generally unknown.
- Some changes of  $F_{k,l}$  are not significant because the linear combination of  $H_{k,l}$  and  $S_{k,l}$  is too sensitive to data variation.

In this paper, fuzzy logic is used to obtain a more robust ranking criterion of input variables. This new criterion is built according to the following fuzzy rules extracted from the human knowledge on the physical meaning of these two elements  $S_{k,l}$  and  $H_{k,l}$ .

If  $S_{k,l}$  is big and  $H_{k,l}$  is big, then  $F_{k,l}$  is big.

If  $S_{k,l}$  is big and  $H_{k,l}$  is medium, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is big and  $H_{k,l}$  is small, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is medium and  $H_{k,l}$  is big, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is medium and  $H_{k,l}$  is medium, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is medium and  $H_{k,l}$  is small, then  $F_{k,l}$  is small.

If  $S_{k,l}$  is small and  $H_{k,l}$  is big, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is small and  $H_{k,l}$  is medium, then  $F_{k,l}$  is medium.

If  $S_{k,l}$  is small and  $H_{k,l}$  is small, then  $F_{k,l}$  is small.

According to these knowledge based fuzzy rules, it can be seen that the coherence degree between human knowledge and measured data plays a more important role in the ranking of variables than the data sensitivity criterion.

These fuzzy rules permit to build a fuzzy model in which  $S_{k,l}$  and  $H_{k,l}$  are taken as two input variables and  $F_{k,l}$  as an output variable. After the fuzzification

procedure, each of them is transformed into a fuzzy variable with three fuzzy values: big, medium and small. For these three variables, we adopt triangular membership functions for the following reasons:

- For a specific product family, the ranges of  $S_{k,l}$ ,  $H_{k,l}$  and  $F_{k,l}$  can be approximately determined from experiments. These ranges permit to determine for each variable, three core numerical values corresponding to their linguistic values "big" (right extreme), "small" (left extreme) and "medium" (medium of the range).
- For each of these three variables, their core values should not be overlapped between them. For example, if one variable is absolutely "medium", then the membership degrees for "small" and "big" should be both 0.
- To simplify, for any value between "small" and "medium" and between "medium" and "big", the corresponding membership degree can be considered as a linear combination between its two extreme cases.

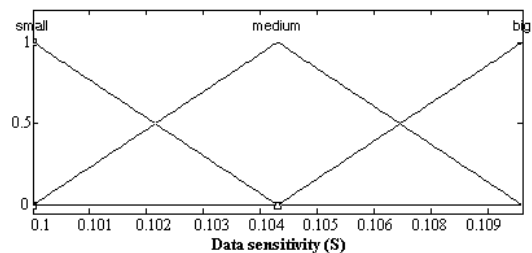


Fig. 4. Membership functions of  $S_{k,l}$

The Mamdani method<sup>14</sup> is used for calculating the output value from input values.

We consider that the output variable  $F_{k,l}$  varies in the range of  $[0,1]$ . The more the value of  $F_{k,l}$  is close to 1, the more the corresponding variable  $x_k$  is relevant. The membership functions of  $S_{k,l}$  are shown in Fig. 4 and can be denoted as  $Triangle(a,a,b)$ ,  $Triangle(a,b,c)$  and  $Triangle(b,c,c)$ . The corresponding parameters  $a$ ,  $b$ ,  $c$  are defined by

$$a = \min_k \{S_{k,l}\}, \quad c = \max_k \{S_{k,l}\} \quad \text{and} \quad b = \frac{a+c}{2}$$

The membership functions of  $H_{k,l}$  and  $F_{k,l}$  are defined in the same way.

This fuzzy logic based selection criterion takes into account both the conformity of human knowledge on process technology and the sensitivity of measured data to the functional properties. Moreover, it is a robust criterion and less sensitive to measuring noises than linear combinations. It is closer to human knowledge on the corresponding process and products. According to this procedure, the most relevant structural parameters of nonwoven products are obtained and will be used in the further modeling procedure. In this way, the complexity of the model can be largely decreased and the parameters of the model will be more concise and easier to be interpreted physically.

### 3. Modeling with Artificial Neural Networks

The artificial neural networks (ANNs) have been used for modeling complex nonlinear problems including various fiber product-related applications.<sup>15, 16</sup>

The existing work on neural network design for small training sets of high dimension has been studied in Ref. 17, 18 and 19. In Ref.17, a diffusion-neural-network has been developed for learning from a small number of samples. In this learning procedure, a number of derived samples are generated from original samples using the method of information distribution.<sup>20</sup> According to Ref.19, modeling with few learning data can be solved by

- reducing the number of input variables by feature selection;
- deploying a rapid, greedy algorithm to identify a good number of nodes to use in hidden layer configuration.

The feature selection problem has been discussed in Section 2. In order to find an optimal neural network architecture, the projection pursuit regression<sup>21</sup> combined with slicing inverse regression<sup>22</sup> is used. Projection pursuit regression introduces a regression family that can be modeled as a sum of subnets of a single hidden layer neural network. This idea can be used for determining the number of nodes in a single hidden layer.

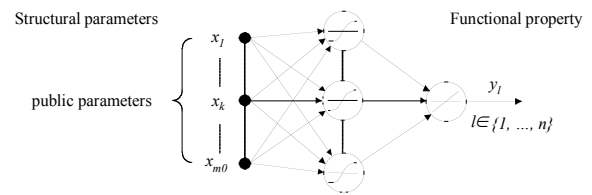


Fig. 5a. General model including only public structural parameters

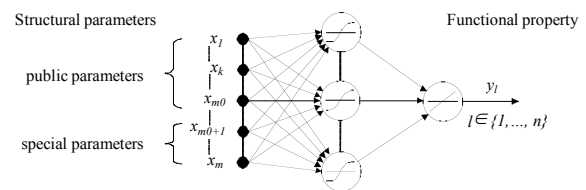


Fig. 5b. Special model including public and special structural parameters

In this section, we use ANNs for modeling the relationship between the structural parameters and the properties of nonwoven fabrics. In general, different nonwoven materials have different structural parameters determined by specific applications, applied technologies and production conditions. Even in the same application field, different technologies are used to manufacture nonwoven products. In this case, the structure of materials varies with applied technology and the corresponding nonwoven products are then classified into a number of families each corresponding to one type of structure. Consequently, all the structural parameters are divided into two groups. One group includes public structural parameters available for all the families of products and the other group includes special structural parameters available for each specific family. Accordingly, two neural network models are built. The general model (Fig. 5a) takes all the public structural parameters as its input variables. This general model can be used by all the families of products. For each specific family, a special model is developed (Fig. 5b). It takes both the public and the special structural parameters of this family as its input variables.

In order to solve the problems related to the lack of available learning data or samples, small scaled ANN models are built. In practice, the performance of an ANN model is strongly related to the relationship between the number of its input variables  $m$  and the total number of learning data  $w$ . According to Ref. 16,

Philippe Vroman, Ludovic Koehl, Xianyi Zeng, Ting Chen

for a three layered neural network used in this paper, this relationship can be described as follows.

The neural network includes  $m$  input variables ( $m$  neurons) each corresponding to one structural parameter. The output layer includes only one neuron corresponding to the specific functional property of interest. The number of hidden neurons, denoted as  $m_2$ , should be at least 2. In this case, we have  $m \times m_2$  weights connecting the input layer to the hidden layer and  $m_2$  weights connecting the hidden layer to the output layer. Also, we have  $m_2 + 1$  biases in this neural network (hidden layer and output layer). Then, the total number of unknown variables in the model (weights and biases) is  $m_2(m+2)+1$ . In order to maintain that the number of unknown variables is no bigger than the number of learning data or constrained conditions, we should have

$$m_2(m+2)+1 \leq w \quad (6)$$

Given a fixed number of learning data  $w$  and considering a minimum of 2 hidden neurons, the maximal value of the number of input variables should be  $(w-5)/2$ .

In our experiments, only 18 learning data are available. Then, we take five input variables in the general model and six input variables in each special model.

In the general model, the transfer functions of the hidden neurons and the output neuron are the hyperbolic tangent function and pure linear function, respectively. The Levenberg-Maquardt fast learning procedure,<sup>3</sup> based on a second order error back propagation algorithm, is then used for determining the parameters of the neural network from the public learning data sets.

In the special model of each family, the weights and biases connecting the public inputs to the hidden layer neurons, as well as those connecting the hidden layer to the output layer, are kept as the same values as in the general model. Only the weights connecting the special input neurons to the hidden layer neurons are adjusted using the error back propagation algorithm.

#### 4. One Industrial Application

In our work, 18 samples describing 3 nonwoven families (6 samples per family) have been used for studying the following functional properties: water

permeability, filtration level, breaking resistance in both the Machine Direction (MD) and the Cross Direction (CD), elongation at peak (MD/CD) and bursting strength. The number of samples is rather limited because of their high production cost and long production time. In fact, these 18 samples are rather diversified and representative because their corresponding process parameters can cover almost all the important working points.

These three nonwoven families are different in the formation (drylaid or spunlaid webs) and the bonding technologies (thermal or chemical bonding). For simplicity, only the modeling procedure and the results related to the water permeability are discussed in this section.

##### 4.1. Identification and selection of the structural parameters

In general, the nonwoven structural parameters are firstly listed according to their own characteristics (raw material, fiber count, crimp and length, thickness, basis weight, porosity, basis weight uniformity, fiber orientation, etc.) and then selected by nonwoven experts according to their possible influence on each functional property. Some structural parameters are difficult to be obtained due to the lack of characterization techniques or measuring instruments.

In our study, 24 public structural parameters are selected by the experts for all the families. They are the fiber length, fiber count, total pore volume, basis weight uniformity, thickness, fiber density, basis weight and so on. The special structural parameters of these three families are the binder rate, the spunbond and the calendaring surface ratio, respectively.

If we take all these 24 structural parameters as input variables, the corresponding model can not be efficient because the number of learning samples is limited to 18, which is too small related to 24 input variables. In that case, the number of parameters for the ANN model would increase to 53 according to the formula (4), considering only 2 hidden neurons.

In order to reduce the complexity of the model and effectively learn the parameters of the model from a small number of data, we select the most relevant



structural parameters using three methods, i.e. the data sensitivity based ranking method, the linear combination based ranking method and the fuzzy logic based ranking method. In the linear combination ranking, the same weight (0.5) is assigned to each of the human knowledge criterion ( $H$ ) and the data sensitivity ( $S$ ). Table 2 gives the most relevant structural parameters related to the water permeability, obtained using these three ranking methods respectively.

Table 2. Ranking of the structural parameters according to their relevancy to water permeability.

Structural parameters	Data sensitivity		Coherence degree $H$	Ranking by linear combination		Fuzzy logic based ranking	
	$S$	rank		$F1$	rank	$F2$	rank
Basis weight	0.1000	7	0.6177	0.3588	1	0.5000	3
Thickness	0.1006	6	0.5212	0.3109	2	0.5000	3
Fiber density	0.1086	1	0.4574	0.2830	3	0.5574	1
Total pore volume	0.1036	5	0.4248	0.2642	4	0.5000	3
Basis weight uniformity	0.1052	2	0.3333	0.2193	5	0.5020	2
Fiber count	0.1052	2	0	0.0526	6	0.3287	6
Fiber length	0.1045	4	0	0.0523	7	0.2119	7

Table 2 shows that the fuzzy logic based ranking ( $F2$ ) takes the fiber density, the basis weight uniformity and the total pore volume as the most relevant structural parameters for the property of water permeability. This result completely conforms to the knowledge of nonwoven experts on the process and the products and is more efficient than the ranking results of  $S$  and  $F1$ . Also, Table 2 denotes a significant difference between the result of the data sensitivity criterion ( $S$ ) and those of the combination of human knowledge and data sensitivity ( $F1$  and  $F2$ ). The result of data sensitivity does not give high ranking orders to total pore volume, thickness and basis weight but emphasizes the fiber count, which is not significant in physical knowledge. In the result of the linear combination based ranking  $F1$ , basis weight uniformity is considered as an irrelevant parameter, which is quite different from the physical knowledge.

In general cases, we can conclude that when using small sets of data for selecting relevant variables, human knowledge seems to positively affect the final ranking

result and more relevant than data sensitivity based criterion. Moreover, the use of the fuzzy logic based ranking brings more efficient results. As this procedure aims at finding a suitable compromise between the data sensitivity and the human knowledge, it gives better results in the combination of the two sources of information.

#### 4.2. Modeling the FP/SP relations for each product family

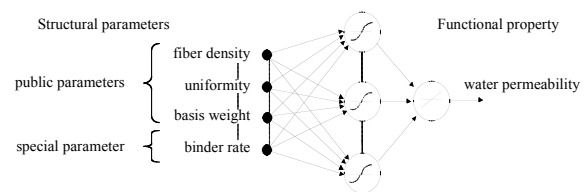


Fig. 6. Special model concerning water permeability for the chemical bonded family

A general model is built using a neural network for all the nonwoven samples. It characterizes the relationship between the selected structural parameters and the corresponding functional property. For the property of water permeability, using the fuzzy logic based ranking method, we take the five most relevant structural parameters as public input variables.

A special model is built for the family of nonwoven materials produced using a specific bonding technology (chemical bonding). Its architecture and parameters are built based on the corresponding general model. The binding rate is added to the set of the input variables of the general model. Fig. 6 shows the special model built for predicting the water permeability with five public parameters (basis weight, thickness, fiber density, total pore volume and basis weight uniformity) as input variables. The special structural parameter (binder rate) is then added to the set of these five input variables.

In this application, the general model is based on 18 samples of three product families. Each family is composed of 6 samples. We use the leave one out technique to test the effectiveness of the general model and the special model. This technique is described as follows. We carry out 18 tests. In each test, we remove one sample from the learning base for testing the models. The remaining 17 samples are used for learning

Philippe Vroman, Ludovic Koehl, Xianyi Zeng, Ting Chen

the general model and the remaining 5 samples of the corresponding family are used for learning the special model. Next, for the testing sample, we calculate the difference between the real value of the water permeability, the output of the general model and the output of the special model. This procedure is repeated for 18 times so that all samples can be removed from the learning base for testing the models.

According to Eq. (6), the number of learning data related to 5 input variables should be no smaller than 15. Then, the leave one out technique with 17 learning data permits to obtain efficient results in the general model.

#### 4.3. Prediction assessment of the product functional properties

Table 3. Experimental results on the functional property “water permeability” for three product families.

Removed sample	Experimental value (l/m <sup>2</sup> /s)	General model		Special model	
		Predicted value (l/m <sup>2</sup> /s)	Error (%)	Predicted value (l/m <sup>2</sup> /s)	Error (%)
1	1103	1004	9.0%	1093	0.9%
2	972	997	2.6%	985	1.3%
3	889	832	6.4%	905	1.8%
4	735	692	5.9%	723	1.6%
5	723	676	6.5%	709	1.9%
6	721	775	7.5%	708	1.8%
7	609	583	4.3%	598	1.8%
8	607	629	3.6%	618	1.8%
9	562	532	5.3%	571	1.6%
10	539	502	6.9%	529	1.9%
11	445	410	7.9%	453	1.8%
12	427	394	7.7%	419	1.9%
13	374	405	8.3%	369	1.3%
14	324	301	7.1%	318	1.9%
15	318	288	9.4%	324	1.9%
16	221	204	7.7%	219	0.9%
17	217	226	4.1%	219	0.9%
18	181	167	7.7%	184	1.7%
Average prediction error (%):		6.55%		1.60%	

Table 3 gives the details of the experimental results on the water permeability and the corresponding predicted results obtained from the general model and the special model. Fig. 7 compares the predicted and the experimental values of the water permeability obtained from the general and the special models and the real physical measures, respectively.

In this experiment, we found that the proposed ANN models give satisfying results with low values of the averaged prediction error, despite the restricted amount of data.

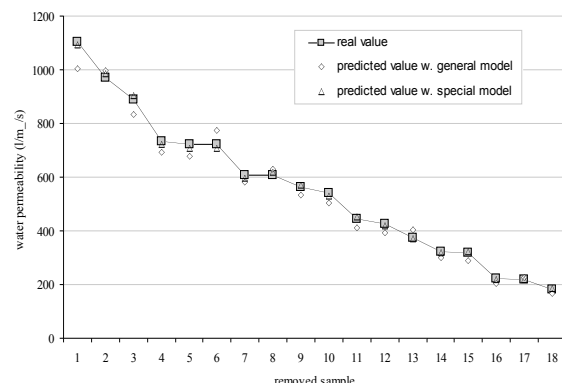


Fig. 7. Comparison general/special models, related to water permeability

From these experimental results, we can see that the special model gives lower prediction errors (averaged error: 1.60%) than the general model (averaged error: 6.55%). This observation can be explained as follows:

1. The general model makes use of samples from several families which differ from each other in many aspects while the special model only uses samples from the same family. The specificity of each family can not be taken into account in the general model.
2. The special model is built based on the same structure as the general model. Only the weights connecting the specific input to hidden neurons are introduced. So, it takes into account both the specificity of each product family and the generality of all families.

#### 5. Conclusions

In this paper, a support system is proposed for optimizing the design of nonwoven products, in accordance with the specifications. The relationship between structural parameters and functional properties of nonwoven products is modeled using artificial neural networks. In order to reduce the complexity of the models and solve the difficulty of insufficient available data, the most relevant structural parameters are selected according to data sensitivity and human knowledge conformity. A fuzzy logic based selection criterion is developed in order to find the best compromise between these two sources of information. The selection procedure of structural parameters allows designers to focus on the most relevant parameters in order to conduct production experiments related to the new product. In the modeling procedure, two models are

defined. A general model is firstly developed for all families of products. It is built from the set of public input variables. A special model is then built for each family of products by adding special structural parameters to the set of public input variables. The simulation of these models allows designers to optimize structure of materials and minimize the number of experiments.

The proposed method has been successfully applied to the nonwoven industry to predict three functional properties: water permeability, breaking resistance and elongation at peak in machine direction. The simulation results show low prediction errors for both the general and the special models. In our future work, this modeling procedure can be further improved by finding new methods to process the existing constraints such as small amount of data, interdependencies between structural parameters and integrate human knowledge on processes and products.

## References

1. Western Europe: nonwoven production +8%. *Technical Textiles* 46(3) (2003) E100.
2. P. Tsai, Theoretical and experimental investigation on the relationship between the Nonwoven structure and the web properties, *International Nonwoven Journal*, Winter (2002) 33-36.
3. D.W. Patterson, *Artificial Neural Networks – Theory and Applications* (Prentice Hall, Singapore, 1996).
4. A.L. Blum and P. Langley, Selection of relevant features and examples in machine learning, *Artificial Intelligence* 97 (1997) 245-271.
5. J.M. Pena, J.A. Lozano, P. Larranaga and I. Inza, Dimensionality reduction in unsupervised learning of conditional Gaussian networks, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 23(6) (2001) 590-603.
6. J.R. Quinlan, Induction of decision trees, *Machine Learning* 1(1) (1986) 81-106.
7. C. Cardie, Using decision trees to improve case-based learning, in *Proc. 11<sup>th</sup> International Conference on Machine Learning* (Amherst, MA, 2003), pp.25-32.
8. R. Battiti, Using mutual information for selecting features in supervised neural net learning, *IEEE Transaction on Neural Networks* 5(4) (1994) 537-550.
9. R. Thawonmas and S. Abe, A novel approach to feature selection based on analysis of class regions, *IEEE Transaction on Systems, Man and Cybernetics - Part B: Cybernetics* 27(2) (1997) 196-207.
10. M. Last, A. Kandel and O. Maimon, Information-theoretic algorithm for feature selection, *Pattern Recognition Letters* 22 (2001) 799-811.
11. R. Kohavi and G.H. John, Wrappers for feature subset selection, *Artificial Intelligence* 97 (1987) 273-324.
12. M. Devaney and A. Ram, Efficient feature selection in conceptual clustering, in *Proc. 14<sup>th</sup> International Conference on Machine Learning*, Morgan Kaufmann, Inc. (San Francisco, CA, 1997), pp.92-97.
13. F. Marcelloni, Feature selection based on a modified C-means algorithm with supervision, *Information Sciences* 151(5) (2003) 201-226.
14. E.H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies* 7 (1975) 1-13.
15. Y. Chen, T. Zhao and B.J. Collier, Prediction of fabric end-use using a neural network technique, *Journal of the Textile Institute* 92(2) (2001) 157-163.
16. J. Fan and L. Hunter, A worsted fabric expert system. II. An artificial neural network model for predicting the properties of worsted fabrics, *Textile Research Journal* 68(10) (1998) 763-771.
17. C. Huang and C. Moraga, A diffusion-neural-network for learning from small samples, *International Journal of Approximate Reasoning*, 35(2) (2004) 137-161.
18. S.J. Raudys and A.K. Jain, Small sample size effects in statistical pattern recognition: recommendations for practitioners, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 13(3) (1991) 252-263.
19. J.L. Yuan and T.L. Fine, Neural-network design for small training sets of high dimension, *IEEE Transaction on Neural Networks* 9(2) (1998) 266-280.
20. C. Huang, An application of calculated fuzzy risk, *Information Sciences* 142(1-4) (2002) 37-56.
21. J.H. Friedman and W. Stuetzle, Projection pursuit regression, *Journal of the American Statistical Association* 76(376) (1981) 817-823.
22. K.C. Li, Slicing inverse regression for dimension reduction, *Journal of the American Statistical Association* 86(414) (1991) 316-327.