Copyright © 2011 Tech Science Press CMES, vol.74, no.3, pp.161-181, 2011

Classification and Optimization Model of Mesoporous Carbons Pore Structure and Adsorption Properties Based on Support Vector Machine

Zhen Yang[1], **Xingsheng Gu**[2] **and Xiaoyi Liang**[1,3]

Abstract: Mesoporous carbons are synthesized by organic–organic self-assembly of triblock copolymer F127 and a new type of carbon precursor as resorcinol–furfural oligomers. Some factors will impact the mesoporous carbons pore structure and properties were studied. The main factors, such as the ratio of triblock copolymer F127 and oligomers, degree of polymerizstry of resorcinol–furfural oligomers, the ratio of resorcinol–furfural oligomers - F/R, and their mutual relations were identified. Aimed at balancing the complex characteristic of mesoporous structure and adsorption properties, a classification and optimization model based on support vector machine is developed. The optimal operation conditions of Barret-Joyner-Halenda (BJH) adsorption cumulative volume and average pore diameter are determined by genetic algorithm support vector classification (GA-SVC). Verification results find that GA-SVC provides an effective method to control and optimize operation conditions and is a new promising theoretical method for material design.

Keywords: Mesoporous carbons; pore structure; adsorption properties; genetic algorithm; support vector machine; classification; optimization

1 Introduction

Mesoporous carbons have attracted considerable attention in the last decade because of their promising applications in such fields as gas storage, catalyst supports and electrode materials[Tamai and Kojima (1997) ;Ryoo, Joo and Krok (2001)]. The synthesis and structure design of mesoporous carbon materials represent an

[1] State Key Laboratory of Chemical Engineering, East China University of Science and Technology; Key Laboratory for Special functional Polymer Materials and Their Related Technologies, Ministry of Education. Shanghai 200237, PR China.

[2] Information Science Institute,East China University of Science and Technology. Shanghai 200237, PR China.

[3] Corresponding author e-mail: xyliang73@sina.com

exciting direction in the research fields of materials science, chemistry, physics and life science, due to their high specific surface areas and confined meso-space for the interaction with encountered matter as well as their good conductivity, anticorrosion, chemical stability and biocompatibility. Lots of novel and efficient synthetic pathways have been developed to fabricate mesoporous carbons materials, giving unique structure characteristics and various potential applications [Lee, Kimand and Hyeon(2006); Chmiola, Yushin, Dash and Gogotsi(2006); Gryglewicz, Machnikowski and Lorenc(2005); Hulicova,Yamashita and Soneda (2005); Salitra, Soffer and Eliad(2000)].

In this work, mesoporous carbons were synthesized by organic–organic self-assembly of triblock copolymer F127 and resorcinol–furfural oligomers. According to the complexity and non-linearity of mesoporous carbons preparation system, typically these studies require a number of experiments to land upon a feasible and desirable solution. Such traditional material design method would be bound to waste lots of manpower and resources. However, due to subjectivity of researchers excessively stressed; it is very difficult to get rational analyzing conclusions. Data redundancy with less information phenomenon is appeared because there are so many incertitude and incompletion data during data processing. Fortunately, theoretical modeling offers a reasonable alternative with which part or total of complex and time-consuming experiments can be replaced. Theoretical modeling is the balance and unification of the experiment induction and deductive reasoning in explore study of the material engineering[Setsuo(1987); Mamalis, Grabchenko and Fedorovich(2001); Aymerich and Serra (2006); Wu,Chiu and Wang (2008)]. Nowadays, intelligent theory and method have been used to predict, estimate and optimize in material engineering[Marmey, Porté and Baquey (2003);Haque and Sudhakar(2002); Arumugam and Rao (2005)]. Pattern recognition is one of the most efficient methods in the theoretical modeling [Kerh,Lai,Gunaratnam and Saunders(2008); Fayyad and stolorz (1997); Chen, Lu and Chen (1999)] which provides new research ways for materials optimized design.

Recently, support vector machines (SVM) has been introduced to solve machine 1earning tasks such as pattern recognition, regression and estimation£®Support vector machine shows advantages in dealing with scared samples, nonlinear and high dimensions. Guo, Wang and Yang (2002) used the SVM to nickel- hydride battery materials feature selection. Based on SVM and leave-one feature selection methods, optimal conditions are founded out. The results showed that forecast effect is totally superior to the other models by using third-order polynomial kernel function in SVM. So, SVM is an efficient way to realize complexity and non-linearity system.

In this paper, tentative theoretical studis for high dimensional complex system were

being investigated. By a new data mining technology-GA support vector machine pattern recognition method, we try to determine predominant factors which influence adsorption properties of mesoporous carbon and extract main parameters of mesoporous structure and adsorption capacity. Aimed at balancing the complex characteristic of mesoporous structure and adsorption properties, the optimal production conditions are found out by model equation. The results show that the GA support vector machine pattern recognition method can not only improve the material selection process, but also can help to obtain more reasonable feature selection system. This paper tends to analyze and decode the connotation of the mesoporous carbon and adsorption properties so as to have a better understanding of mesoporous carbon system. It is hopeful for providing an effective theoretical basis and achieving breakthrough results in material design.

2 Experiment and Data Processing

2.1 Experiment

Triblock copolymer Pluronic F127 (EO106PO70EO106, Mw=12,600) were purchased from Sigma-Aldrich Corporation. Resorcinol ($\geq$99.5%), furfural($\geq$99.5%), hydrochloric acid (36.0-38.0%) ($\geq$ 99.0%), six methylene-4-amine, and ethanol were purchased from Shanghai Sinopharm Chemical Reagent Co.Ltd.

In this experiment, the effect of the mesoporous carbon is investigated by altering the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F / R. The procedures were as follows:

Mesoporous carbons are synthesized by means of solvent EISA method using resorcinol and furfural as carbon sources and triblock copolymer pluronic F127 as pore-directing agent. Typically, 1.75g resorcinol is dissolved in 45g ethanol with stirring at 298K followed by addition of 2.25g furfural and 0.5g 0.01mol/L hydrochloride ethanol solutions. In addition, reaction time is changed from 15 to 60 min. The mixture is stirred for 30 min to obtain resorcinol-furfural oligomer solution. Then 0.2g hexamethylenetetramine and 1-6g F127 is added with stirring for 10 min to obtain a clear dark red solution. The solution is transferred into smooth glass plates to evaporate off ethanol at room temperature for about 6h, and the resulting films are heated in an oven at 393K for 24h. The samples are then carbonized at 1173K for 3h to decompose triblock copolymers and to obtain mesoporous carbons [Long, Qiao, Zhan, Liang and Ling(2009)].

2.2 Experimental data

The main data obtained by altering the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F / R are listed in Table 1 [Long (2009)].

2.3 Two-dimensional Analysis of Experimental Data

To determine internal and external factors that affect mesoporous carbon properties, two-dimensional analysis of experimental data are restricted to fix a number of influence factors and change a certain factor such as the ratio of F127/RF. Analysis results can be obtained simply by altering the ratio of F127/RF or degree of polymerization of R/F or the ratio of F/ R. However, we can only obtain independent impact for mesoporous carbon properties by this method. Also, the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F / R are often interrelated. So, how to get total effects factors on the mesoporous carbon properties is more meaningful.

2.3.1 Relationships among the ratio of F127/RF, N_2 adsorption–desorption and pore size distribution

In same reaction conditions, a series of different structure of mesoporous carbon are prepared by altering the ratio of F127/RF. Results indicate that total pore volume and mesoporous pore volume increase with the increase ratio of F127/RF within certain range. Also, we find that the porous structure are changed from the microporous to the disorder mesoporous, three-dimensional cube, two-dimensional hexagonal, and finally to worm-like structure with the increase ratios of F127/RF. The most common causes of structures'change are due to different degree of swelling of the EO chain.

2.3.2 Relationships among the ratio of F/R, N_2 adsorption–desorption and pore size distribution

When the ratio of F127/RF is 0.75 and reaction time is 30min, along with increase of the ratio of F/R from 0.8 to 2.5, surface area of samples have been decreased from 810 m^2/g to 677m^2/g and total pore volume decreased from 0.6 cm^3 / g to 0.43 cm^3/g, and average pore diameter have been decreased from 3.8 nm to 3.4 nm. Results indicate that the total pore volume and mesoporous pore volume increase with the increase ratio of F127/RF. The decreased pore structure parameter may result from the decreased degree of ordering.

2.3.3 The relationship between the degree of polymerization of Resorcinol- furfural oligomers, N_2 adsorption–desorption and pore size distribution

As seen form the Table 1, the BET surface area, total pore volume and BJH adsorption decrease distinctly with the increase of degree of polymerization of Resorcinol-furfural oligomers.

Microstructure of samples is observed by TEM technology. Under the ratio of

Table 1: Main parameters of mesoporous carbon obtained by altering the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F/R

NO.	F127/RF[a]	F/R[b]	Time[c]	S^{d}_{BET}	V^{e}_{t}	V^{f}_{mes}	D^{g}_{BJH}
	wt./wt.	mol/mol	min	m^2/g	cm^3/g	cm^3/g	nm
1	0.2	1.5	30	450	0.24	0.03	3.2
2	0.3	1.5	30	524	0.34	0.18	3.3
3	0.4	1.5	30	677	0.43	0.21	3.7
4	0.5	1.5	30	785	0.49	0.25	3.6
5	0.5	1.5	45	668	0.4	0.24	3.7
6	0.5	1.5	60	481	0.26	0.12	3.4
7	0.55	2.5	30	604	0.32	0.11	3.7
8	0.6	1.5	30	719	0.5	0.24	3.5
9	0.75	0.8	30	810	0.6	0.42	3.8
10	0.75	1.5	30	770	0.53	0.32	3.7
11	0.75	1.5	45	671	0.46	0.32	3.7
12	1	1.5	30	736	0.55	0.35	3.7
13	1.25	1.5	30	803	0.61	0.42	3.8
14	0.75	1.5	60	537	0.35	0.14	3.6
15	0.75	2.5	30	677	0.43	0.2	3.4
16	1.36	0.8	30	657	0.91	0.74	7.4
17	1.5	1.5	30	805	0.68	0.47	3.9

[a] the mass ratio of F127/R+F;
[b] the mole ratio of furfural/resorcinol;
[c] reaction time of resorcinol with furfural before the addition of hexamethylenetetramine;
[d] BET surface area;
[e] total pore volume;
[f] BJH adsorption cumulative volume;
[g] BJH adsorption average pore diameter.

F127/RF=0.5 and reaction time being 30, 45 and 60min, three-dimensional cube, two-dimensional hexagonal and disordered mesoporous carbon structure can be obtained, respectively. Results indicate that, with the same compounds, the porous structure change from the micro-porous to the disorder mesoporous, three-dimensional cube, two-dimensional hexagonal, and finally to disordered structure with the increase of polymerization reaction time [Grosso,Cagnol and Soler(2004)].

Mesoporous materials possess variety of components and complex feature. Actually, its adsorption property and pore structure are governed by combination factors, furthermore, the ratio of F127/RF, degree of polymerization of resorcinol/furfural and the ratio of F/R are conjoined yet opposed factors are varied greatly. It is difficult to express the relationships among main factors affecting the performance by a precise mathematical model. The factors that control and affect the performance of mesoporous materials are random and fuzzy variables, however, two-dimensional experimental data analysis method restrict to fix a number of influence factors and change a certain factor such as the ratios of F127/RF, which might lead to get inaccurate or wrong analysis results.

As powerful data mining method, artificial intelligence is good at searching for mathematical laws underlying limited data. For example, pattern recognition is suitably used to solve non-linear classification and discrimination problems [Kowalski and Bender (1972), Wold (1997)]. In this paper, GA support vector machines pattern recognition method is used to classify and optimize adsorption properties and pore structure for mesoporous materials.

3 Genetic Algorithm Support Vector Machine Classification

3.1 Support Vector Classification

SVM was developed by Vapnik V and his colleagues at AT&T Bell Laboratories in 1995 [Guo, Wang and Yang (2002)]. Unlike most traditional neural network models based on the principle of empirical risk minimization, SVM is motivated by statistical learning theory based on the principle of structural risk minimization. It can be applied in classification and regression. Most of intelligent theory is based on law of large number which represents a statistical law, however, the number of samples is always limited in material design, and if too much data are used to build intelligent models, the noise will add to the model and lead to reduce predictive ability. SVM will help to overcome most of above problems because of its standout learning capability, and it has been one of the hot points in machine learning research.

SVM technique has been proposed in references[Cristinini and Taylor(2000); Couran and Hilbert (1953)], where more details about the algorithmic implementation can be found. In this part, we only give a description of support vector classification relative to this paper.

SVM method is proposed by linear optimal separating surface of linear separable problem. From statistical learning theory, linear optimal separating surface is the separating hyperplane and has the best generalization performance. Optimal separating surface, not only can correctly classify all the training samples, but also can

get the largest distance from the nearest point of training samples to classification surface (see Fig. 1) [Boser, Guyon and Vapnik (1992);Nello and John (2000)].

In Fig. 1, L corresponding to the largest interval hyper-plane is called as optimal partition hyperplane. The distance of the two parallel dotted lines and is the half of the largest interval. Maximum intervals and the optimal hyper-plane only are determined completely by the sample points which fall to the border, and they are called as support vector, such as the sample points in Fig. 1.

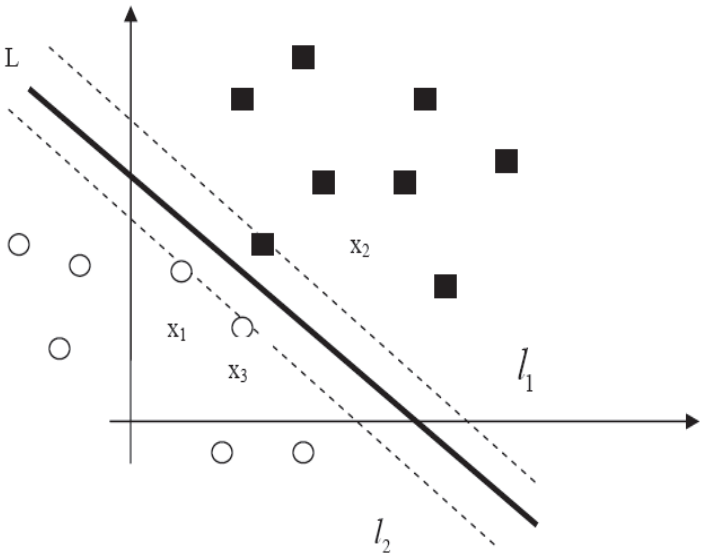

Figure 1: Schematic diagram of optimal separating hyperplane

The diagram of support vector classification networks is shown in Fig. 2.

In Fig. 2, training set's vectors are separated to two classes (x_1, y_1), ... , (x_n, y_n), x∈R, y ∈(-1, +1). The hyperplane can be written as $\omega^T x + b = 0$.

If the training data are linearly separable, then:

$$y_i(\omega^T x_i + b) - 1 \geq 0, i = 1, 2, \cdots, n \tag{1}$$

$$min\left|\omega^T x_i + b\right| = 1, i = 1, 2, \cdots, n \tag{2}$$

The decision rule can be written as:

$$f_{\omega,b}(x) = sgn(\omega^T x + b) \tag{3}$$

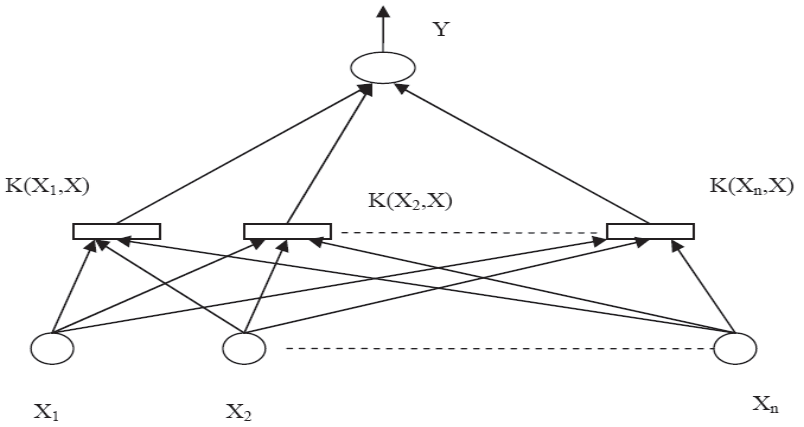

Figure 2: The diagram of support vector classification networks

In Eq. (3), ω denotes the weight variable, and b is the bias.

The optimization is now a quadratic programming problem:

$$\begin{aligned} &minJ(\omega) = \tfrac{1}{2}(\omega^T \omega) \\ &s.t : y_i(\omega^T x_i + b) \geq 1, i = 1, 2, \cdots n \end{aligned} \tag{4}$$

This problem has a global optimum. Lagrangian transformed the above optimal separating surface into its dual problem, namely:

$$L(\omega, b, \Lambda) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^{n} \lambda_i [y_i(\omega^T x_i + b) - 1] \tag{5}$$

In Eq. (5), $\Lambda = \{\lambda_1, \lambda_2 \cdots \lambda_n\}$ are the Lagrange multipliers. Hence:

$$F(\Lambda) = \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j \tag{6}$$

Lagrange multipliers are non-zero only if $y_i(\omega^T x_i + b) = 1$. For these cases, vectors are called as support vectors. Then the optimal separating hyper plane is given by $\omega^* = \sum_{i=1}^{n} \lambda_i * x_i y_i$ and bias is given by $b^* = \frac{1}{2}(\omega^*)^T (x_r + x_s)$.

where, x_r and x_s are support vectors when y_r=1 and y_s=-1.

The classifier is quantified as $f_{\omega,b}(x) = sgn(\omega^T x + b)$.

For linearly no separate problem, we should add relaxation term $\xi \geq 0$ to Eq. (2), and the optimization problem can be written as:

$$\begin{aligned} &minJ(\omega,\xi) = \tfrac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}\xi_i \\ &s.t: y_i(\omega^T x_i + b) \geq 1-\xi_i, i = 1,2,\cdots n, \xi \geq 0 \end{aligned} \tag{7}$$

In Eq. (7), C is error penalty parameter to control the penalty degree of classifying error. In fact, the dual problem of optimal classification surface in linearly no separate problem is almost same as that of linear separable problem, but the condition changes to$0 \leq \alpha_i \leq C, i = 1,2,\cdots n$.

SVC can not only solve linear classification problem but also can solve nonlinear classification problems by nuclear function (the inner product function) technology. By nonlinear mapping, SVM transfers input space into high-dimensional space and finds out an optimal classification surface (the best feature space) in the transformed space. In feature space, kernel function $K(x,x_i) = <\Phi(x)\Phi(x_i)>$is introduced.

Thus, classification function is defined as follow:

$$f(x) = sgn[\sum_{SV}\alpha_i * y_i K(x,x_i) + b*] \tag{8}$$

In transformed feature space, kernel function $K(x,x_i)$ can avoid lots of operation and greatly reduce computation complexity.

3.2 Parameters Optimization of Support Vector Machine Classifier Based on Genetic Algorithm

Keerthi. S S and Lin C J [Keerthi and Lin (2003)] found that radial basis function (RBF) is better than other kernel functions, so we select RBF as the kernel function for support vector machine classifier. RBF is usually represented by:

$$K(x_k,x) = \exp(-\|x-x_k\|^2/\sigma^2). \tag{9}$$

There are two adjustable parameters in RBF SVM classifier, including penalty parameter C and RBF kernel function parameters such asσ^2. How to choose the best kernel parameters is an important problem for SVC because it will directly affect its accuracy and generalization performance. Kernel function parameter σ^2 has effect on complexity in high-dimensional feature space. The penalty parameter C can adjust confidence range of learning machine and risk-experience ratios in feature space. Therefore, to obtain better generalization performance for SVM classifier, we should firstly select suitable σ^2 to map data into feature space, and then select suitable C to adjust confidence range of the learning machine and optimum

risk-experience ratios in feature space. Also, it has great significance to save considerable machining time and improve operation precision in simultaneously optimizing σ^2 and C [Cherkassky and Ma (2004)]. The essence of above problem is optimization. In this paper, genetic algorithm is used to optimize the above two parameters.

GA, based on a direct analogy to Darwinian natural selection and genetics in biological systems, is a promising alternative to conventional heuristic methods. Based on the Darwinian principle of 'survival of the fittest' [Xing and Xie (1999)], GA can obtain the optimal solution after a series of iterative computations. Implementation processes of GA include chromosome design, population generation and fitness study. GA deals with data in large search spaces, hence, has less chance to get local optimal solution than other algorithms. More details about the algorithmic implementation can be found in reference Davis (1991). In this study, the detailed description of steps of chromosome design, fitness function and evolutionary system for GA-based SVC parameters optimization are as follows:

1) Set t=0 where t represents evolutional generation.

2) The ranges of C and σ^2 are as follows:

$$\begin{cases} 0 \le \sigma^2 \le \max(\left\|x_i - x_j\right\|^2 \times 10^{-2}) \\ 0 \le C \le \max(\alpha) \end{cases} . \tag{10}$$

More than this, optimization speed and efficiency are taken into account in our study.

3) Float encoding is used for randomly given initial C and σ^2.

4) Fitness function is defined as:

$$F(\sigma^2, C) = M/\Delta E. \tag{11}$$

5) Optimization objective is evaluated by $\min(|f(x) - y|)$.

Only the above certain condition is met, GA-based SVC can turn to step 10.

6) Set t=t+1.

7) We use roulette selection in each population.

Selection operation is defined as:

$$P_k = F(\sigma^2, C)/\Sigma F(\sigma^2, C), \tag{12}$$

In Eq. (12), P_k is selection probability.

8) Crossover operators are computed by Eq. (13).

$$\begin{cases} Child_1 = factor \times Parent_1 + (1 - factor) \times Parent_2 \\ Child_2 = factor \times Parent_2 + (1 - factor) \times Parent_1 \end{cases} \tag{13}$$

In Eq. (13), factor is a random number between 0 and 1.

9) A random number is generated to determine the variation parameters according to mutation probability and random parameter values are generated to replace the original parameters. Using uniform mutation operation with smaller mutation probability, original value of gene seat is replaced by encoding string.

10) If termination conditions are met, genetic algorithm will be terminated. The best C and σ^2 will be output. On the contrary, steps 4-9 will be repeatedly executed until C and σ^2 are satisfied with model error.

4 Classification results and discussion

Main procedures of support vector machine classifier optimization modeling for mesoporous carbon materials pore structure and properties are as followed:

4.1 Data standardization and normalization

In order to eliminate dimension different, Eq. (14) is used for data standardization and normalization:

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}. \tag{14}$$

All input and output data are standardized and normalized to [0, 1].

4.2 Parameter Optimization

When population size is set as 60; the biggest evolution algebra is set as 200, and hybrid probability and mutation probability are set as 0.85 and 0.15. According to Eq. (7) and Eq. (10), range of σ^2 and C are set as [0,100] and [0,500], respectively. To verify the algorithm's effectiveness proposed in this paper, we use different training and testing sample sets in classification parameter optimization. Table 2 shows classification results comparison of different sample sets.

From Table 2, when the number of training samples is 12 and evolution generations are 58, we can get the optimal solutions: C=311.48, σ^2= 2.36 and classification accuracy rate = 94.18%.

Table 2: Classification results comparison of different sample sets

The numbers of training samples	Accuracy rate	σ^2	C	Evolution generations
8	76.47	51.21	405.48	90
10	82.35	36.08	318.56	86
12	94.18	2.36	311.48	58
16	88.24	5.03	286.15	65
17	82.35	4.92	371.51	73

4.3 Relationships among process parameters, pore structure parameters and adsorption properties

In this study, process parameters include F127/RF ratio, polymerization degree of resorcinol-furfural oligomers and the ratio of F/R. Variables X (composition of polymer chemical material), Y (pore structure of mesoporous carbon) and Z (characteristics value of adsorption) are used in various classifications of parameters. The meanings and markers of specific parameters are shown in Table 3.

Table 3: The meanings and markers of parameters

Material chemical composition		Pore structure		Characteristics value of adsorption	
Variable	Symbol	Variable	Symbol	Variable	Symbol
Raito of F127/RF	X1	BET surface area	Y1	BJH adsorption cumulative volume	Z1
Raito of F / R	X2	Total pore volume	Y2	BJH adsorption average pore diameter	Z2
Degree of R-F oligomers	X3				

4.3.1 Study of relationships among process parameters, pore structure parameters and BJH adsorption cumulative volume using pattern recognition

There are two methods used in this part of work, one is Fisher method (the principle of Fisher method can be found in reference (the principle of Fisher method see reference Duda, Hart and Stork (2003)), the other is GA-SVC algorithm. The data

processing program are provided with self-compiled software and Matlab [Math Works Incorporation (2008)] in Intel Pentium IV computer.

1) Traditional Fisher pattern recognition

In accordance with adsorption properties indicators of the mesoporous carbon, excellent class for BJH adsorption cumulative volume should be reach 0.25cm^3/g or higher. In Table 1, samples which adsorption capacity are greater than or equal to 0.25cm^3/g are classed as "1" grade and marked as "1". Samples which BJH adsorption cumulative volumes are less than 0.25cm^3/g are classed as "2" grade and marked as "2". x_1, x_2, x_3, y_1, y_2 are used to implement a five-dimensional space. Traditional Fisher pattern recognition is used to study the space distribution regularity for BJH adsorption cumulative volume of samples "1" and "2". Fig. 3 and Fig. 4 show that good regular relationship can be obtained by this method.

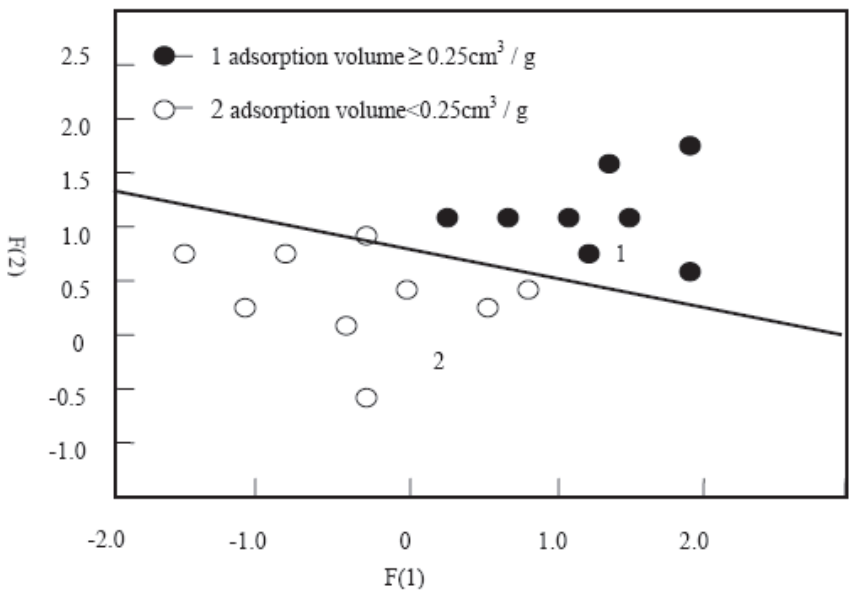

Figure 3: BJH adsorption cumulative volume Fisher classification projection

In Fig. 3:

$$\begin{cases} F(1) = 3.662 \times 10^{-1}[X_1] + 4.763 \times 10^{-1}[X_2] - 1.529 \times 10^{-1}[X_3] \\ +4.437 \times 10^{-1}[Y_1] - 1.11 \times 10^{-1}[Y_2] - 1.375 \\ \\ F(2) = 2.784 \times 10^{-1}[X_1] - 5.648 \times 10^{-1}[X_2] - 0.958 \times 10^{-2}[X_3] \\ +2.236 \times 10^{-1}[Y_1] - 0.724 \times 10^{-1}[Y_2] + 4.908 \end{cases} \tag{15}$$

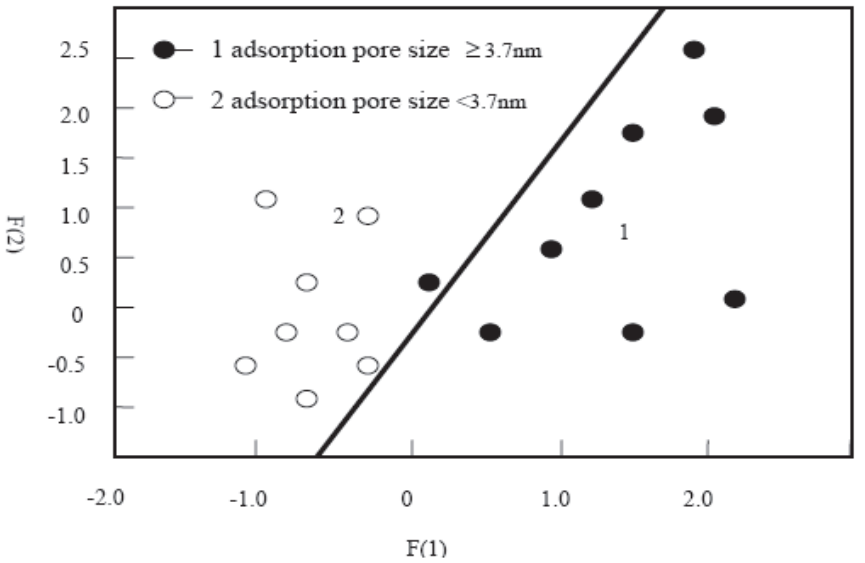

Figure 4: BJH adsorption average pore diameter Fisher classification projection

In Fig. 4:

$$\begin{cases} F(1) = 2.426 \times 10^{-1}[X_1] - 5.214 \times 10^{-1}[X_2] + 1.311 \times 10^{-1}[X_3] \\ +3.459 \times 10^{-1}[Y_1] - 1.198 \times 10^{-1}[Y_2] - 4.203 \\ \\ F(2) = 2.6820 \times 10^{-1}[X_1] + 3.546 \times 10^{-1}[X_2] - 1.142 \times 10^{-2}[X_3] \\ +3.615 \times 10^{-1}[Y_1] - 0.854 \times 10^{-1}[Y_2] + 3.517 \end{cases} \tag{16}$$

2) Genetic Algorithm-Support Vector Classification and optimization method

GA-SVC algorithm is used to establish the classification and optimization model. Table 1 makes a list of the seventeen samples in which twelve of them are used as training samples and the last five are used as test samples to estimate extrapolate predictive ability. Samples according to different BJH adsorption cumulative volume are divided into two classes: samples of which adsorption volume are greater than or equal to 0.25cm^3/g are classed as "1" grade and marked as "1",and the other samples of which adsorption volume are less than 0.25cm^3/g are classed as "2" grade and marked as " 2".

When we set the training sample number as 12, C as 311.48 andσ^2 as 2.36, the data in Table 1 are take into convex quadratic dual programming problem. BET surface area and total pore volume are used as dependent variables and BJH adsorption cumulative volume as output variable.

The calculations obtained by this method are as follows:

$Q(\alpha)_{max}$= 0.0743

So,

$$\omega* = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \alpha_i \alpha_i * y_i y_i * (x_i, x_i *) = (0.3045, 0.0786, 0.0368, 0.2314, 0.1025) \tag{17}$$

From the constraint condition$y_i(\omega^T x_i + b) \geq 1, i = 1, 2, \cdots n$, the classification threshold b is set as -14.2578. As a result, the classification functional expression of BJH adsorption cumulative volume can be written as:

$$\begin{aligned} g(x) &= sgn(f(x)) = sgn\left\{ \sum_{i \in sv} \alpha_i y_i < \Phi(x) \cdot \Phi(x_i) > + b \right\} \\ &= sgn[(0.3551, 0.1657, 0.0849, 0.3483, 0.1134) - 14.2578] \end{aligned} \tag{18}$$

The first twelve samples are used to examine classification function, and the classification results are accordant with actual results in all cases. Results are as follows: $y_{13} = y_{16} = y_{17} = 1, y_{14} = y_{15} = 2$

The results show that the unknown samples 13, 16 and 17 should belong to excellent class and samples 14 and 15 should belong to poor class, which are in complete agreement with the experience of experts [Long,Qiao,Zhan,Liang and Ling(2009)].

3) The comparative study of BJH adsorption cumulative volume model between Fisher and GA-SVC

To make objective evaluation Fisher and GA-SVC method, we select and test twelve training samples and five testing samples by 30 times. Table 4 shows that the erroneous recognition(ER) of these two methods are 8% and 4%, respectively.

Table 4: BJH adsorption cumulative volume classification comparison between Fisher and GA support vector classification

Tested samples					Fisher					ER	GA-support vector classification					ER
1	2	3	4	5	2	2	2	1	1	8%	-1	-1	-1	1	-1	4%
6	7	8	9	10	2	2	2	1	1		-1	-1	-1	1	1	
11	12	13	14	15	1	1	1	2	2		-1	1	1	-1	-1	
13	14	15	16	17	2	2	2	1	1		-1	-1	-1	1	1	
1	3	5	7	9	2	2	2	2	1		-1	-1	-1	-1	1	

4.3.2 *Study of the relationships among process parameters, pore structure parameters and BJH adsorption average pore diameter of mesoporous carbon using pattern recognition*

1) Traditional Fisher pattern recognition

In accordance with BJH adsorption average pore diameter indicators of the mesoporous carbon, excellent class for absorption pore size should be greater than or equal to3.7nm. In Table 1, the samples which adsorption pore size are greater than or equal to 3.7nm are classed as “1” grade and marked as “1”,and the samples which adsorption pore size are less than 3.7nm are classed as “2” grade and marked as “2”. Fisher pattern recognition is used to study the space distribution regularity for BJH adsorption average pore diameter of samples "1" and "2". It can also obtain good regular relationship by Fisher method.

2) Genetic Algorithm -Support Vector Classification and optimization method

GA-SVC algorithm is used to establish classification and optimization model of process parameters, pore structure and BJH adsorption average pore diameter. Table 1 make a list of seventeen study samples in which twelve of them are used in training samples and the last five are used in test samples to estimate extrapolate predictive ability of GA-SVC method.

Samples according to different BJH adsorption average pore diameter are divided into two classes: samples of which adsorption pore size values are greater than or equal to 3.7nm are classed as “1” grade and marked as “1”, and the other samples of which adsorption pore size values are less than 3.7nm are classed as “2” grade and marked as “2”.

The BJH adsorption average pore diameter impact relationship classification function can be written as

$$\begin{aligned} g(x) &= sgn(f(x)) = sgn\left\{\sum_{i\in sv}\alpha_i y_i < \Phi(x)\cdot\Phi(x_i) > +b\right\} \\ &= sgn[(0.5124, 0.3227, 0.1294, 0.4579, 0.1051) - 9.5348] \end{aligned} \tag{19}$$

The BJH adsorption average pore diameter classification results are accordant with actual results and the experience of experts. Results are as follows:

$y_{13} = y_{16} = y_{17} = 1, y_{14} = y_{15} = 2$

3) The comparative study of BJH adsorption average pore diameter model between Fisher and GA-support vector classification

The alternative twelve and five samples are selected as training and testing samples and they are tested by 30 times. The results are shown in Table 5, in which erroneous recognition of BJH adsorption average pore diameter are 12% and 4%, respectively.

Table 5: BJH adsorption average pore diameter classification comparison between Fisher and GA-support vector classification

Test samples					Fisher					ER	GA- support vector classification					ER
1	2	3	4	5	2	2	1	2	1		-1	-1	1	-1	1	
6	7	8	9	10	1	1	1	1	1		1	1	-1	1	1	
11	12	13	14	15	2	1	1	2	2	12%	-1	1	1	-1	-1	4%
13	14	15	16	17	2	2	2	1	1		1	-1	-1	1	1	
1	3	5	7	9	2	1	1	1	1		-1	1	1	1	1	

Our comparisons show that GA-SVC and optimization model has higher correct rate. We can use these results to find and verify that the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F/R have been recognized as three main influence factors of adsorption properties. Such studies achieve the same results as the experts for several reasons. Firstly: with increasing F127/RF ratio, the mesostructures transform from the disordered phase to ordered cubic to hexagonal phases and then to disordered wormlike phase and finally to mesocellular structure. Increasing F127/RF ratio shifts the free energy balance and leads to the development of local composition fluctuations. Next, during the period of reaction, the molecular weight of oligomers gradually increase, and steric hindrance, chain elongation and chain branching gradually increase and lead to a dramatic decrease of hydrophilic properties of oligomer. As the polymerization proceeds, the oligomers link together and grow larger, which can assume fewer mixed configurations than small oligomers. Forthermore, when the ratio of F127/RF is 0.75 and the reaction time is 30min, along with the increase of the ratio of F/R from 0.8 to 2.5, results indicate that, within certain range, the total pore volume and mesoporous volume increase with the increasing of F127/RF ratio. The decreased pore structure parameter may result from the decreased degree of ordering [Melosh,

Lipic and Bates(1999); Urade, Bollmann and Kowalski (2007); Alberius, Frindell and Hayward (2002); Urade, Wei and Tate (2007)]. There are still some other influence factors, but all of them, the ratio of F127/RF, degree of polymerization of resorcinol/furfural and the ratio of F/R are most prominent.

5 Conclusions

The results of our research show that pattern recognition, support vector machines and other data mining methods can be used as effective methods of data processing for complex system such as mesoporous carbon materials. The results also show that pore structure and adsorption properties can be improved by appropriately increasing the ratio of F127/RF, degree of polymerization of resorcinol/furfural (R/F) and the ratio of F/R, which are in line with the previous research in XRD patterns and TEM analysis. Based on this, pattern recognition image processing can be used to identify hidden relationships or connections among properties of mesoporous carbon materials. Using support vector machine algorithm, we can obtain BJH adsorption cumulative volume and average pore diameter classification function and classification and optimization model with low erroneous recognition. This paper develops GA-support vector machines for optimizing C andσ^2 based on analysis of the influence factors on kennel parameters.

The advantages of GA-support vector machines classification and optimization modeling are as follows:

i) Support vector machines classification can achieve high precision no need to have too much samples.

ii) GA based-support vector classification is developed to obtain optimal parameters and avoid blindness of manual selection. By use of this method, it is simple and reliable for determining structure parameters of SVM classifier. After appropriate modifications, the method can be used to parameter optimization design of support vector machine regression and other intelligent methods.

iii) By use of GA based-support vector machines pattern recognition method, excellent adsorption properties model for mesoporous carbon can be obtained, which provides the theoretic study for mesoporous carbon material design and offers some helpful suggestions about good processing conditions for preparation samples.

Hence, GA-SVC classification and optimization model developed in this study can be effectively used for mesoporous carbon and worth to be popularized. But the causal relationship between the dependent variable and target variables of the model needs further research. All of these conclusions indicate the value of our future work.

Acknowledgement: This work was partly supported by National Science and Technology Ministry (2009BAE72B04), National Science Foundation of China (50730003 and 50672025), National Project of Scientific and Technical Supporting Programs Funded by Ministry of Science & Technology of China (2007BAE55B00), National High Technology Research and Development Program of China (2007AA05Z311).

References

Tamai, H.; Kanata, T.; Kojima, S. (1997): Preparation of mesoporous activated carbon fibers and their adsorption properties, *Adv Mater*, vol. 9, pp. 55-58.

Ryoo, R.; Joo, S. H.; Krok. M. (2001): Ordered mesoporous carbons, *Adv Mater*, vol. 13, pp. 677-681.

Alberius, P. C. A.; Frindell, K. L.; Hayward, R. C. (2002): General predictive syntheses of cubic, hexagonal, and lamellar silica and titania mesostructured thin films, *Chem Mater*, vol. 14(8), pp. 3284-3294.

Arumugam, M. S.; Rao, M. V. C. (2005): On the optimal control of single-stage hybrid manufacturing systems via novel and different variants of particle swarm optimization algorithm, *Discrete Dyn Nat Soc*, vol. 19, pp. 257–279.

Aymerich, F.; Serra, M. (2006): An ant colony optimization algorithm for stacking sequence design of composite laminates. *CMES: Computer Modeling in engineering & sciences,* vol. 13, pp. 49-66.

Boser, B.; Guyon, I.; Vapnik, V. (1992): A training algorithm for optimal margin classifiers, Fifth Annual Workshop on Computational Learning Theory.

Chen, N. Y.; Lu, W. C.; Chen, R. L. (1999): Chemometric methods applied to industrial optimization and materials optimal design, *Chemometrics and Intelligent Laboratory System*, vol. 45, pp. 329-333.

Cherkassky, V.; Ma, Y. (2004): Practical selection of SVM parameters and noise estimation for SVM regression, *Neural Networks*, vol. 17(1), pp. 113-126.

Chmiola, J.; Yushin , G.; Dash, R.; Gogotsi , Y. £¨2006£©: Effect of pore size and surface area of carbide derived carbons on specific capacitance, *J. Power Sources*, vol. 158(1), pp. 765-772.

Courant, R.; Hilbert, D. (1953): Method of mathematical physics(Volume I), New York: Springer Verlag.

Cristinini, N.; Taylor, J£®(2000): An Introduction on Support Vector Machine and Other Kernel-based Learning Methods, Cambridge£ºCambridge University Press.

Davis, L. (1991): Handbook of genetic algorithms. New York: Van Nostrand Reinhold.

Duda, R. O.; Hart, P. E.; Stork, D. G£®(2003): Pattern classification, Beijing: Mechanical Industry Press.

Fayyad, U.; Stolorz, P. (1997): Data mining and KDD£ºpromise and challenges, *Future Generation Computer Systems*, vol. 13, pp. 99-115.

Grosso, D.; Cagnol, F.; Soler, I. G. (2004): Fundamentals of mesostructuring through evaporation-induced self-assembly, *Adv Funct Mater*, vol. 14(4) , pp. 309-322.

Gryglewicz, G.; Machnikowski, J.; Lorenc, G. E. (2005): Effect of pore size distribution of coal-based activated carbons on double layer capacitance, *Electrochimca Acta* , vol. 50(5), pp. 1197-1206.

Guo, Z. Wang, Z. X.; Yang, J. (2002): SVM-based feature selection method and applications, *Computer and Applied Chemistry*, vol. 19 (6) , pp. 703-705.

Haque, M. E.; Sudhakar, K. V. (2002): ANN back propagation rediction model for fracture toughness in microalloy steel, *Int J Fatigue*, vol. 24, pp. 1003-1017.

Hulicova, D.; Yamashita, J.; Soneda, Y. (2005): Supercapacitors prepared from melamine-based carbon, *Chem. Mater*, vol. 17(5), pp. 1241-1247.

Keerthi, S. S.; Lin, C. J. (2003): Asymptotic behaviors of support vector machines with Gaussian kernel, *Neural Compu* , vol. 15(3), pp. 1667-1689.

Kerh, T.; Lai, J. S.; Gunaratnam, D.; Saunders, R. (2008): Evaluation of seismic design values in the taiwan building code by using artificial neural network. *CMES: Computer Modeling in engineering & sciences,* Vol. 26, pp. 1-12.

Kowalski, B. R.; Bender, C. F. (1972): Pattern recognition: a powerful approach to interpreting chemical data, *Journal of the American Chemical a Society,* vol. 94, pp. 5632-5639.

Lee, J.; Kim, J.; Hyeon, T. £¨2006£©: Recent progress in the synthesis of porous carbon materials, *Advanced Materials,* vol. 18(16), pp. 2073-2094.

Long, D. H.; Qiao, W. M.; Zhan, L.; Liang, X. Y.; Ling, L. C. (2009): Effect of template and precursor chemistry on pore architectures of triblock copolymer-templated mesoporous carbons, *Microporous and Mesoporous Materials*, vol. 121, pp. 58-66.

Long, D. H. (2009): Synthesis and structure design of mesoporous carbon materials, East China University of Science and Technology.

Mamalis, A. G.; Grabchenko, A. I.; Fedorovich, A. I. (2001): Development of an expert system of diamond grinding of superhard polycrystalline materials considering grinding wheel relief, *Advanced Manufacturing Technology*, vol. 17(7), pp. 498-508.

Marmey, P.; Porté, M. C.; Baquey, C. (2003): PVDF multifilament yarns grafted with polystyrene induced by γ-irradition: Influence of the grafting parameters on the mechanical properties, *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms,* vol. 208, pp. 429-433.

Math Works Incorporation. (2008): MATLAB user manual R2008a, Math Works.

Melosh, N. A.; Lipic, P.; Bates, F. S. (1999): Molecular and macroscopic structures of transparent block copolymer-silica monoliths, *Macromolecules*, vol. 32(13), pp. 4332-4342.

Nello, C.; John, S. T. (2000): An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press.

Salitra, G.; Soffer, A.; Eliad, L. (2000): Carbon electrodes for double-layer capacitors I. relations between ion and pore dimensions, *J. Electrochem. Soc.* vol. 147(7), pp. 2486-2493.

Setsuo, O. (1987): Application of artificial intelligence for material design, *Proceedings of Computer Aided Innovation of New Materials*, vol. 15(3), pp. 263-270.

Urade, V. N.; Bollmann, L.; Kowalski, J. D. (2007): Controlling interfacial curvature in nanoporous silica films formed by evaporation-induced self-assembly from nonionic surfactants. II. Effect of processing parameters on film structure, *Langmuir*, vol. 23(8), pp. 4268-4278.

Urade, V. N.; Wei, T. C.; Tate, M. P. (2007): Nanofabrication of double-gyroid thin films, *Chem Mater*, vol. 19(4), pp. 768-777.

Vapnik, V. N. (1998): Statistical Learning Theory, A Wiley-Interscience Publication, John Wiley and Sons, lnc. USA.

Wold, S. (1997): Chemometrics: theory and application, ACS Symposium Series, No. 52, Washington D . C.

Wu, C. P.; Chiu, K. H.; Wang, Y. M. (2008): A differential reproducing kernel particle method for the analysis of multilayered elastic and piezoelectric plates. *CMES: Computer Modeling in engineering & sciences,* vol. 27, pp. 163-186.

Xing, W. S.; Xie, J. X. (1999): Modern heuristic algorithm. Beijing: Tsinghua University Press.