Novel QSAR Combination Forecast Model for Insect Repellent Coupling Support Vector Regression and *K*-Nearest-Neighbor

[1, 2] LI-FENG WANG, [3]XIAN-SHENG TAN, [1, 2]ZHE-MING YUAN* AND [2, 3]LIAN-YANG BAI**
[1]*Hunan Provincial Key Laboratory of Crop Germplasm Innovation and Utilization, Changsha, China;*
[2]*College of Bio-safety Science and Technology, Hunan Agricultural University, Changsha, China;*
[3]*Hunan Institute of Humanities; Science and Technology, Loudi, China.*
zhmyuan@sina.com*, bailianyang2005@yahoo.com.cn**

(Received on 3rd September 2012, accepted in revised form 19th December 2012)

Summary : To improve the precision of quantitative structure-activity relationship (QSAR) modeling for aromatic carboxylic acid derivatives insect repellent, a novel nonlinear combination forecast model was proposed integrating support vector regression (SVR) and K-nearest neighbor (KNN): Firstly, search optimal kernel function and nonlinearly select molecular descriptors by the rule of minimum MSE value using SVR. Secondly, illuminate the effects of all descriptors on biological activity by "multi-round enforcement resistance-selection". Thirdly, construct the sub-models with predicted values of different KNN. Then, get the optimal kernel and corresponding retained sub-models through subtle selection. Finally, make prediction with leave-one-out (LOO) method in the basis of reserved sub-models. Compared with previous widely used models, our work shows significant improvement in modeling performance, which demonstrates the superiority of the present combination forecast model.

Key Words: QSAR; Insect repellent; Combination forecast; SVR; KNN.

Introduction

Repellent is a chemical substance, which acts on insects and disturbs their natural behavior. As a main measure of personal protection, insect repellent has the advantages of small dosage, low price, no environmental pollution etc. Therefore, it has broad prospects in preventing ectoparasite and insect diseases [1-3].

Quantitative structure -activity relationship (QSAR) has become and will continue to be an important tool for drug synthesis and discovery, it can be established to research the regular relationship between molecular structural descriptors and biological activity. Traditional modeling methods of QSAR based on empirical risk minimization, such as multiple linear regression (MLR), stepwise linear regression (SLR), principal component regression (PCR), partial least square regression (PLS) and artificial neural networks (ANN), have many defects [4-6] . Support vector machine (SVM) is a new kind of learning machine based on statistical learning theory presented by Vapnik, which uses structural risk minimization instead of empirical risk minimization [7]. It can solve the problems of small-sample, non-linear, over-fit, high dimension, etc. SVM representing a group of supervised learning techniques, which first applied in pattern recognition, has been successfully used to solve both classification and regression problems [8-11].

In present study, we proposed a nonlinear screening method to select descriptors based on support vector regression (SVR), a novel SVR combination forecasting method to make prediction in basis of *K*-nearest neighbor (KNN) and then, applied this approach in QSAR research on aromatic carboxylic acid derivatives insect repellent. Compared with reference models, the performance of our novel approach is very encouraging.

Results and Discussion

Optimal Kernel Function and Retained Descriptors

With all 35 samples, the molecular descriptors were screened by SVR model with LOO method, the obtained MSE values are shown in Table-1. From which, we can draw that polynomial kernel function t=1, d=2 is the optimal kernel function with corresponding MSE value of 0.031231. For optimal kernel function, 2 descriptors are sequentially rejected and 6 descriptors relating to repellent activity are retained. Hence, the polynomial kernel function t=1, d=2 and the reserved 6 descriptors are finally applied to subsequent global prediction and KNN prediction.

Table-1: MSE of different kernel functions with retained descriptors

Kernel Function	t=0	t=1,d=2	t=1,d=3	t=2	t=3
Retained Descriptors	2,3,6,7,8	1,4,5,6,7,8	1,2, 4,5,7	3,5,7,8	2,4,5,6,7,8
MSE	0.070303	0.031231	0.040613	0.04556	0.071219

*: 1= B, 2= LogP, 3= $LogP^2$, 4=σ^o, 5= MR1, 6= MR2, 7=I, 8=MV.

In order to investigate the relative effects of descriptors on repellent activity, compulsive elimination was conducted to select all descriptors with minimum MSE as principle. The concrete data are shown in Table-2, from which the relative importance of each descriptor to repellent predictive activity could be listed as follows: σ^o **>** MR1 **>** I **>** MW **>** LogP **>** MR2 **>** $LogP^2$ **>** B. It indicates that descriptor σ^o makes the largest contribution to repellent activity is from and descriptor B has the least contribution.

Global Prediction and KNN Prediction

Because the sample set is small in this case, the *K*-value was selected uniformly as *K*=1 (corresponding to nearest-neighbor predictions), *K*=7, 14, 21, 28 (corresponding to different KNN prediction), and *K*=34 (corresponding to global prediction). The predict repellent activities and corresponding MSE values calculated with LOO method based on optimal kernel function t **=** 1, d=2 and 6 retained descriptors.

Combination Forecast

Although global prediction is slightly superior to the near-neighbor predictions setted in this work, actually, the better *K*-value of near neighbors might also exists. We cannot foreknow the best *K*-value prior to application, but the search of optimal *K*-value from training set is notsuitalbe. On the one hand it will cause a great amount of calculation, on the other hand the optimal *K*-value of each tested sample is obviously different and there exists no commonly applicative optimal *K*-value. Combination forecast always has better prediction accuracy and stability [12]. Hence we adopt the combination forecasting to predict repellent activity with near neighbor *k* vales of 1, 7, 14, 21, 28 and 34.

The 6 sub-models are respectively constructed on the basis of different *K*-value using kernel function t=1, d=2 and then, the actual activities and predictive values of sub-model compose a new "combination forecasting sample set" (35×7matrix). The further kernel function optimization and sub-model screen were carried out for this new "combination forecasting sample set" and the results are that optimal kernel function is determined as t=3 and the retained sub-models are 21 and 34 near neighbors models (Table-3).

Using optimal kernel function t=3, we applied LOO method to take final combination forecast with predicted values of 21 and 34 nearest neighbor sub-models as new descriptors. The forecasting activity values and relevant MSE values are shown in Table-4, from which we can draw that combination forecasting made better performances than KNN prediction, nearest neighbor prediction and global prediction. Even though the value of *K* and the number of sub models are changed at random, the performance of combination model is still superior in prediction accuracy and stability.

The predict results of model MLR (SPSS13.0, Enter Method) and SLR (SPSS13.0, Stepwise Method) and fitting results of model SLR obtained with LOO method with 8 descriptors were also shown in Table-4, from which we can conclude that the performance of novel combination model (MSE=0.043, MAPE=16.878) is much better than that of traditional linear model MLR (MSE=0.247, MAPE=32.097) and SLR (MSE=0.137, MAPE=26.875), close to fitting results of SLR (MSE=0.031, MAPE=14.631). Compared with Hansch free energy analysis (MSE=0.1432 0.031, MAPE=28.29414.631) adopted in reference [2], our novel model also shows significant superiority.

Experimental

Data Set

The dataset used in this study was taken from the work of Xue *et al.* [2], 8 molecular structure descriptors of 35 aromatic carboxylic acid derivatives insect repellents and their actual repellent activity against housefly are shown in Table-5. The used descriptors include boiling point (B), hydrophobic parameter (LogP, $LogP^2$), electrical parameter (σ^o), steric parameter (MR1, MR2), indicator variable (I) and molecular weight (MW).

There had been 40 data samples in literature [2], but 5 incomplete data samples were rejected due to the lack of individual descriptors in this study.

Table-2: Enforcement optimization of descriptors based MSE.

Descriptor	B	LogP	$LogP^2$	σ^o	MR1	MR2	I	MW	Rejected Descriptor
1st round	0.0715	0.0907	0.0892	0.0845	0.0890	0.0844	0.0848	0.0865	B
2nd round	--	0.0757	0.0697	0.0724	0.0845	0.0719	0.0824	0.0705	$LogP^2$
3rd round	--	0.0523	--	0.0767	0.0728	0.0338	0.0740	0.0539	MR2
4th round	--	0.0400	--	0.0713	0.0637	--	0.0781	0.0669	Log p
5th round	--	--	--	0.0525	0.0466	--	0.0680	0.0446	MW
6th round	--	--	--	0.0905	0.1219	--	0.0816	--	I
7th round	--	--	--	0.0705	0.0519	--	--	--	MR1

Table-3: MSE of different kernel functions with retained descriptors of sub-models.

Kernel Function	t=0	t=1,d=2	t=1,d=3	t=2	t=3
Retained Sub-models	*K*=1,14,21,34	*K*=1,7,28,34	*K*=7,14,21,28,34	*K*=7,28,34	*K*=21,34
MSE	0.058436	0.065570	0.078589	0.048381	0.044793

Table-4: Predict and fit value of different models with LOO method.

Actual Activity	Fit Value of SLR	Predicted Value of MLR	Predicted Value of SLR	Predicted Value of *K*=1	Predicted Value of *K*=7	Predicted Value of *K*=14	Predicted Value of *K*=21	Predicted Value of *K*=28	Predicted Value of *K*=34	Combination Forecast Value
1.83	2.181	2.278	2.278	2.180	2.430	1.651	1.810	1.852	1.785	1.850
2.18	2.428	2.487	2.556	2.510	2.289	2.217	2.110	2.190	2.179	2.208
2.43	2.656	2.683	2.683	2.550	2.436	2.377	2.467	2.437	2.462	2.899
2.66	2.827	2.914	2.914	2.890	2.806	3.088	3.074	2.980	2.954	2.355
2.83	2.747	0.876	2.738	2.810	2.647	2.763	2.663	2.840	2.755	2.662
2.75	2.924	3.934	2.877	2.880	2.810	2.848	2.896	3.027	3.256	3.151
2.92	3.800	2.394	2.569	3.150	2.893	2.870	2.847	2.896	3.046	3.295
3.80	1.858	1.606	1.606	1.830	2.080	2.152	2.066	2.230	2.266	2.169
1.86	2.276	2.235	2.235	2.550	2.516	2.446	2.434	2.418	2.435	2.710
2.28	2.501	2.492	2.492	2.530	2.528	2.566	2.575	2.589	2.684	2.730
2.50	2.770	2.771	2.736	2.550	2.530	2.434	2.494	2.485	2.508	2.611
2.77	2.676	2.641	2.641	2.780	2.746	2.996	2.717	3.012	3.038	2.993
2.68	2.772	2.736	2.736	2.600	2.838	2.255	2.581	2.625	2.863	3.116
2.77	2.277	2.699	3.058	2.550	2.718	2.439	2.792	2.498	2.027	2.033
2.28	2.577	2.549	2.932	2.880	2.865	3.341	3.098	2.914	2.824	2.863
2.58	2.293	2.356	2.356	2.520	2.395	2.761	2.041	2.108	2.180	2.170
2.29	2.554	2.565	2.565	2.180	2.430	1.651	1.810	1.852	1.785	2.253
2.55	2.822	2.905	2.905	2.510	2.289	2.217	2.110	2.190	2.179	2.390
2.82	2.939	2.818	2.895	2.550	2.436	2.377	2.467	2.437	2.462	3.260
2.94	2.472	2.380	2.380	2.890	2.806	3.088	3.074	2.980	2.954	2.946
2.47	2.512	2.509	2.509	2.560	2.520	2.470	2.496	2.621	2.500	2.285
2.51	2.783	2.759	2.759	2.520	2.927	2.806	2.876	2.929	2.909	2.785
2.78	3.000	2.945	2.945	3.220	2.820	2.638	2.457	2.678	2.618	2.975
3.00	3.048	2.994	2.994	2.560	3.281	2.912	2.693	3.015	2.544	3.285
3.05	2.899	2.933	2.933	2.860	2.520	2.553	2.410	2.505	2.467	3.134
2.90	3.195	3.391	3.391	3.190	2.808	2.875	2.830	2.859	2.911	3.145
3.19	2.764	2.822	2.822	3.220	2.933	2.942	3.108	3.101	3.164	2.948
2.76	2.818	2.805	2.805	3.190	2.875	2.908	3.581	2.933	3.211	2.952
2.82	2.987	2.929	2.929	2.910	2.810	2.850	2.620	3.004	2.742	3.176
2.99	3.130	3.113	3.113	2.640	2.897	2.860	3.053	2.954	2.953	2.952
3.13	3.215	3.159	3.159	2.910	2.861	2.862	2.846	2.228	2.602	3.137
3.21	3.263	3.299	3.299	3.190	3.269	3.130	3.149	3.033	3.015	2.937
3.26	3.348	3.391	3.391	3.220	3.084	3.108	3.060	3.089	3.132	3.053
3.35	2.963	3.094	3.150	3.190	3.119	3.065	3.005	3.153	3.257	2.593
2.96	3.177	3.136	3.136	3.220	3.187	3.053	2.964	2.982	3.120	3.230
MAPE	14.631	32.097	26.875	27.257	24.601	32.294	33.609	28.773	27.704	16.878
MSE	0.031	0.247	0.137	0.070	0.104	0.118	0.117	0.094	0.050	0.043

Table-5: Molecular structure descriptors and actual repellent activity of insect repellent.

Compounds	B	LogP	LogP^2	σ^o	MR1	MR2	I	MW	MV
B1	199	2.11	4.45	0	12.87	1.03	0	136.15	1.83
B2	213	2.65	7.02	0	16.35	1.03	0	150.18	2.18
B3	219	3.06	9.36	0	21.03	1.03	0	164.21	2.53
B4	249	3.73	13.91	0	25.64	1.03	0	178.22	2.55
B5	263	2.21	4.48	0	19.81	1.03	1	149.19	2.78
B6	281	2.71	7.34	0	28.04	1.03	1	177.25	3.15
B8	397	2.89	8.35	0	30.46	1.06	1	189.26	3.8
P1	215	1.83	3.35	0	16.48	1.03	0	150.18	2.18
P2	226	2.37	5.62	0	21.31	1.03	0	164.21	2.51
P3	245	2.78	7.73	0	25.99	1.03	0	178.23	2.55
P4	262	3.45	11.9	0	30.6	1.03	0	192.26	2.89
P5	288	2.37	5.62	0	23.56	1.03	1	163.22	2.81
P6	297	2.44	5.95	0	32.95	1.03	1	191.27	2.88
P7	332	2.84	8.07	0	36.37	1.03	0	212.25	2.03
P8	312	2.67	7.13	0	33.21	1.03	1	203.25	2.6
S1	223	2.12	4.49	-0.09	12.87	2.85	0	152.15	2.18
S2	232	2.66	7.07	-0.09	16.35	2.85	0	166.18	2.52
S3	241	3.19	10.18	-0.09	21.03	2.85	0	180.2	2.56
S4	271	3.74	13.99	-0.09	25.64	2.85	0	194.23	3.19
S7	298	3.01	9.06	-0.09	30.15	2.85	0	214.22	2.63
M1	248	2.24	5.02	0	12.87	7.87	0	166.18	2.52
M2	247	2.78	7.73	0	16.35	7.87	0	180.2	2.86
M3	251	3.19	10.18	0	21.03	7.87	0	194.23	3.19
M4	267	3.36	11.29	0	25.64	7.87	0	208.26	3.22
M5	331	2.26	5.11	0	19.81	7.87	1	179.22	2.86
M6	278	2.5	6.25	0	28.04	7.87	1	207.27	2.91
M8	334	2.26	5.11	0	30.46	7.87	1	219.28	2.64
E1	252	2.78	7.73	0.11	12.87	12.47	0	180.2	2.85
E2	255	3.32	11.02	0.11	16.35	12.47	0	194.23	3.19
E3	261	3.73	13.91	0.11	21.03	12.47	0	208.26	3.22
E4	273	4.4	19.36	0.11	25.64	12.47	0	222.28	3.25
E5	315	2.8	7.84	0.11	19.81	12.47	1	193.25	3.19
E6	298	2.91	8.47	0.11	28.04	12.47	1	221.3	3.25
E7	299	3.97	15.76	0.11	30.15	12.47	0	242.27	2.68
E8	312	2.86	8.18	0.11	30.46	12.47	1	233.31	3.27

Principle and Software of SVR

In SVR, the basic idea is to map the input samples into a higher-dimensional feature space and then make linear regression in this space. The transformation is implemented by proper kernel function $K(x_i, x_j)$. Assume that the input samples are p-dimension vector, n samples and output values expressed as $(x_1, y1)$. $(x_n,y_n)\in R^p\times R$, then the learning problem of support vector regression becomes a quadratic programming problem. The ε -insensitive loss function is usually adopted to express the loss. If the allowed error is designed as ε and the error of the sample x is ξ and then, the loss is ignored when $|\xi|\le\varepsilon$, or else the loss is designated as $|\xi|-\varepsilon$. The regression function is defined as $f(x)=\sum_{i=1}^{n}(\alpha_i-\alpha_i^*)K(x,x_i)+b$, where a and a^* are solved p-dimension vector, $K(x_i,x_j)=\varphi(x_i)\cdot\varphi(x_j)$ is kernel function expressed as dot product of two $\varphi(x)$, and $\varphi(x)$ is the mapped image in higher-dimensional feature of point x. It is can be seen from analysis that the selection of kernel function has important effect on the predictability of SVR, and the usage of kernel function makes it possible to solve the nonlinear regression without definite mapped image of $\varphi(x)$. In $f(x)$, if the weight value (a-a^*) unequal to zero, then the sample x_i is called support vector. Obviously, the number of support vectors determines the computation complexity and has strong correlation with prediction accuracy [7, 13-15].

LIBSVM2.86 from http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html was adopted to implement SVR model in this article. The process was completed by LIBSVM2.86 with C++ compiled by us and this self-compiling program was tested and verified by successive verification.

KNN

KNN is an analogy-based algorithm, which has been widely used in classification problems. Its basic idea is to find out K nearest points with tested sample in multidimensional space and judge the category of tested sample according to the ones of those K points. Similar to classification, KNN can be also used in regression problems, that is, calculating the tested point with the K nearest points as training samples. Obviously, the value range must be of $1\le K\le n-1$ in regression problems and its two ends

correspond to the nearest neighbor prediction and global prediction, respectively. The standard Euclidean distance *Dist* (x_i, x_j) is used as KNN distance function in regression. Assume two samples are X_i= {$X_{i,1}$, $X_{i,2}$ $X_{i,m}$} and X_{j} = {$X_{j,1}$, $X_{j,2}$ $X_{j,m}$} then the Euclidean distance can be denoted as following equation:

$$Dist\left(x_i, x_j\right) = \sqrt{\sum_{a=1}^{m}\left(x_{i,a} - x_{j,a}\right)^2} \qquad (1)$$

In SVR, the selection of training samples for prediction with leave-one-out (LOO) method are based on the global, namely all samples except tested one participate in training process. There often exists heterogeneity in sample set, so the precision of global prediction is not always the optimal. For example, if the samples were divided into two categories, then the training set belonging to the second category will be interfered with the prediction results of tested sample which belongs to the first category. In this case, the KNN prediction with a appropriate *K*-value maybe much better than global prediction. The *K*-value could be estimated through system cluster, nonlinear map, etc. But it is very difficult and time-consuming to search for the best *K*-value for each tested sample [14, 15].

Construction of Combination Forecast Model Based on SVR and KNN

Combination forecast may have better performance than individual methods [16]. Therefore, a novel combination forecast model was approached based on SVR and KNN with different *K*-value in this study.

Optimization of Kernel Function

Kernel function plays an important role in SVR prediction. For five given functions (liner kernel function t=0, polynomial kernel function t=1, d=2, polynomial kernel function t=1, d=3, rbf kernel function t=2, sigmoid kernel function t=3), the optimal kernel function could be selected through making predictions with LOO method and comparing their corresponding mean squared error (MSE) values. LOO method is an objective and strict forecasting performance test method which used in the field of machine learning. It can be used to determine how accurately a learning algorithm predicts the tested data which was not trained on [17].

Nonlinear Screen of Molecular Descriptors Based on SVR

The screen of descriptors is an important step in QSAR analysis. A good correlation between the selected descriptors and the bioactivity implies better bioactivity predictions [18]. This study adapted the nonlinear descriptor selection strategy in the basis of SVR. Assume there are *n* samples, *p* descriptors, based on MSE value, indistinctive descriptors for improving prediction precision are successively swept with multi-round last-elimination method from the SVR model which contains all descriptors. For first-round selection , $MSE_{(X1,\ X2,\ Xi,\ xp)}$ is mean square error with *p* input variables and $MSE_{(x1,x2,\ xi-1,\ xi+1\ xp)}$ is mean square error with the *ith* input variable deleted. If $MSE_{(x1,\ x2,\ xi-1,\ xi+1\ xp)} < MSE_{(x1,\ x2,\ xi-1,\ xi=1\ xp)}$, it indicates that the *ith* input descriptor has very poor correlation with bioactivity. So we remove it out and carry out the next round selection (now change the *p* into *p-1* in the formulae). Contrariwise , there is no descriptor to reject, and the elimination is over.

The optimal kernel function and retained descriptors are applied to construct the sub-models.

Construction of Sub-Model

From optimal kernel function and retained descriptors, the Euclidean distance between tested sample and the rest *n-1* samples is calculated respectively; according to different *k*-value, different near neighbors are used as training samples to make prediction with LOO method based on SVR and then, the obtained predictive value comprises one sub-model. We can select 3 to 10 *K*-values uniformly in the range [1, n-1] to construct 3 to 10 sub-models (the bioactivity predicted value of sub-models in combination forecast samples equals to molecules descriptors in original samples).

Kernel Optimization and Sub-Model Selection for Combination Forecast Samples

Based on minimum MSE, referencing to the sections "*Optimization of kernel function*" and "*Nonlinear screen of descriptors based on SVR*", the optimal kernel function and retained sub-models can be obtained by carrying out kernel optimization and sub-model selection with LOO method.

Combination Forecast

Make prediction based on optimal kernel function and retained sub-models with LOO method.

Prediction Evaluation Index and Algorithm Program

MSE and mean absolute percentage error (MAPE) are applied to evaluate the performance of models:

$$MSE = \frac{1}{n}\sum (y - \hat{y})^2 \quad (2)$$

$$MAPE(\%) = \frac{1}{n}\sum \frac{|y - \hat{y}|}{y} \times 100 \quad (3)$$

where y represents actual repellent activity value, $\hat{y}$ is fitting or forecasting repellent activity value and n denotes sample number.

Conclusions

When a strict LOO prediction was conducted, the tested sample could not involve in the selections of kernel function, descriptors and sub-models, etc. However, after repeat trials, the optimal kernel function, reserved descriptors and sub-models are all demonstrated to be similar whether the number of samples is being 35 or 34. Hence, the LOO tested samples in each prediction model all belong to independent ones in this work.

The novel SVR-KNN combination model avoids the puzzle to select the optimal K value which single KNN model will meet, and conquers the drawbacks of less useful information and weak stability for 1NN. And also the novel model produces better results with predictive ability compared with other methods. It was proved to be a useful tool in prediction of drugs activity. This model has potential to be widely used in QSAR and other relative fields. Furthermore, the nonlinear relationship adopted in present model can describe the relationship between the structural parameters and the bioactivity of aromatic carboxylic acid derivatives insect repellent, which provides the clues for further optimization of present repellents.

Acknowledgments

This work was supported by a grant from the Science Foundation for Distinguished Young Scholars of Hunan Province, China (10JJ1005), the Research Fund for the Doctoral Program of Higher Education of China (No. 20124320110002), and Open Subject of Hunan Provincial Key Laboratory for Biology and Control of Plant Disease and Insect Pests (BCH201101).

References

1. M. Kirk, *Encyclopedia of Chemical Technology Supplement*, John wiley and Sons, New York, p. 12 (1984).
2. F. Q. Xue, H. Q. Wang, R. C. Zhao, J. Yang, Y. C. Wang, Z. Z. Xu, X. L. Du and J. Y. Zhang *Scientia Agricultura Sinica*, **30**, 77 (1997).
3. S. Alexander, S. Andrea, W. Christian, Klampfl and B. Wolfgang, *Microchim Acta*, **148**, 151(2004).
4. B. Jayendra, A. K. Bhonsle, R. K. Bhattacharjee and Gupta, *Journal of Molecular Modeling*, **13**, 179 (2007).
5. R. Kunal and P. R. Partha, *European Journal of Medicinal Chemistry*, **44**, 2913 (2009).
6. H. F. Mohammad and G. Sajjad, *Bioorganic and Medicinal Chemistry*, **15**, 7746 (2007).
7. N. Vapnik, *the Nature of Statistical Learning Theory*, Spring-Verlag Press, New York, p. 112 (1995).
8. W. M. Zhong, D. Y. Pi and Y. X. Sun, *Journal of Central South University of Technology*, **12**, 591 (2005).
9. Q. Wu, W. Y. Liu and Y. H. Yang, *Journal of Central South University of Technology*, **14**, 442 (2007).
10. K. C. Chi, G. B. Kingsman and H. Wong, *Annals of Operations Research*, **87**, 129 (1999).
11. P. Alexei and K. Mikhail, *Stochastic Environmental Research and Risk Assessment*, **22**, 647 (2008).
12. K. C. Chi, B. G. Kingsman and H. Wong, *Annals of Operations Research*, **87**, 129 (1999).
13. J. C. Meng and L. Xia, *International Journal of Infrared and Millimeter Waves*, **28**, 413 (2007).
14. X. Y. Wang and Z. O. Wang, *Forecasting*, **12**, 24 (1988).
15. G. X. Chai and Y. Q. Wu, *Acta Mathematicae Applicatae Sinica*, **29**, 542 (2006).
16. Y. C. Li, B. Li and T. J. Fang, *Information and Control*, **33**, 303 (2004).
17. Z. X. Yu, C. B. Zhou and J. P. Li, *Chinese Journal of Rock Mechanics and Engineering*, **24**, 2468 (2005).
18. G. Shovanlal, D. Bikash and S. T. J. Soma, *Bioorganic and Medicinal Chemistry*, **12**, 1493 (2004).