CHEMICAL ENGINEERING TRANSACTIONS

VOL. 51, 2016

A publication of AIDIC

The Italian Association
of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Tichun Wang, Hongyang Zhang, Lei Tian

ISBN 978-88-95608-43-3; **ISSN** 2283-9216

DOI: 10.3303/CET1651172

Forecasting of Mine Discharge Based on Phase Space Reconstruction and SimpleMKL

Xin Zhang[a,b], Yinghai Guo*[a], Youkuo Chen[a,b]

[a]School of Resources and Earth Sciences China University of Mining and Technology Xuzhou, Jiangsu 221116, China.
[b]Yankuang Group Company Limited, Zoucheng, Shandong 273500, China.
teasta@163.com

Forecasting of mine discharge is crucial to water disaster prevention in mines. The hydrogeological is complex, caused by internal and exterior factors in a mine that makes the forecasting much more difficult. The paper puts forward a kind of prediction method of mine discharge based on Phase Space Reconstruction and SimpleMKL. By analysis and studying the hydrogeology of the study field, while getting the maximal Lyapunov exponent $\lambda = 0.1844$ >0 by wolf algorithm, we know mine discharge has Chaos characteristics and can find out the hidden evolution rule of mine discharge time series by Phase Space Reconstruction, and as SimpleMKL input and output parameters, so as to transform the nonlinear problem into linear solution in multi-kernels mapping space. After choosing mine discharge date from August 2004 to February 2005 to validate, the results show that this method is feasible and possesses a high precision and very strong practical value.

1. Introduction

The reason that causes water bursting is that the natural balance of the rocks is destroyed during the process of mine tunnelling or face mining, surrounding water, under the action of hydrostatic pressure and mine ground pressure, access into face via faults, aquiclude and weak spots inside the strata(]Zhang C.W.et.al.(2004)).The mine groundwater system is gradually evolving with mining depth and mining area gradually increasing, the process of system evolution show by the original steady state to the unsteady state, and then reach another steady state under the system self-adjusting. The system evolution process is like an iterative process, each initial conditions of the iteration is different, mainly displays in the difference of groundwater quantity and ground water level. Due to the comprehensive effect of geologic structure, formation lithology, hydrometeorology and any other reasons, each initial condition shows a relative independence and randomness(Tang and Xiao, 2006), therefore mine discharge time series has chaos characteristics. Chaos Phenomena is unique because of its completely definite, simple nonlinear system model, if we can find the internal law of nonlinear system and establish its model, then forecasting of its uncertainty is possible. Takens' Embedded Theorem provides theoretical bases for prediction of chaotic time series(Takens, 1981),many nonlinear prediction methods that forecasting chaotic time series were proposed(Meng and Peng, 2007), Adaptive Nonlinear Filtering Prediction Method(Zhang and Xiao, 2000), The Prediction Method Based on Neural Network(Gu and Zhai, 1995), for instance, Feed forward Neural Network Nonlinear Prediction Method based on Phase Space Reconstruction is able to accurately predict the low dimensional chaotic sequence, but due to minimized sample point error of feed-forward neural network model in training(Cui et al., 2005), 'over-fitting' phenomenon is inevitable, which will lead to the model poor generalization ability.

Support Vector Machine(SVM) now become a new hotspot after neural network, it was established on the basis of VC dimensional theory and structural risk minimization principle of Statistical Theory(Vapnik, 1999), SVM has been seen as a good replacement of traditional learning methods, especially in smallsamples, high dimensional nonlinear case, and has good generalization performance(Chapelle et al., 1999).

Whereas, SVM is single kernel methods both based on individual character space. Different kernel functions have different characteristics, so installed performance of a kernel function varies greatly in different applications, and the construction or choice of kernel functions has no completely theoretical basis, at present, commonly used kernel functions are linearly kernel function, polynomials kernel function, Radial Basis

Please cite this article as: Zhang X., Guo Y.H., Chen Y.K., 2016, Forecasting of mine discharge based on phase space reconstruction and simplemkl, Chemical Engineering Transactions, 51, 1027-1032 DOI:10.3303/CET1651172

Function (RBF) kernel function, Sigmoid kernel function and Gauss kernel function etc.(Genton, 2001), in addition, when sample characteristics contain Heterogeneous Information(Bach F. R.(2008)), large sample size(Rakotomamonjy et al., 2008), multi-dimensional unnormalised data(Kingsbury et al., 2005) or non-flat data in high dimension space distribution characteristics(Yang et al., 2006), to deal with all samples through single kernel function mapping is not reasonable. To address these issues, a great deal of research about the method of kernel combination are come out, namely, multiple kernel learning(Bach, 2004).
SimpleMKL can be extended to other SVM algorithms with only little changes such as regression, clustering, singular detection, mutli-classification etc. SimpleMKL converges in fewer times than other MKL algorithms.
The Phase Space Reconstruction based on chaotic dynamic system are similar to machine learing, the article attempt to approach a method of prediction mine discharge based on Phase Space Reconstruction and simpleMKL, so SimpleMKL theoretical research can be applied to practical work, thus, serveing for the safety production of mines better.

2. Methodology

2.1 Phase Space Reconstruction based on chaotic time series

Phase Space Reconstruction is also called system reconstruction, original Phase Space Structure can be excavated out based on one-dimensional time series, so the intrinsic nature can be found out. At present, the commonly used method is time delay. The method is approached by Packard et al.(Packard N.H.et al.(1980)), Takens consolidated the mathematics basics. Any weight is decided by others of the system, so the interrelated vector information is included in any weight.
Supposed that the time series: $x=\{x_i=1,2,\ldots,N\}$, then any vector in the Phase Space Reconstruction follows:

$$X = \{X_i \mid X_i = [x_i, x_{i+t}, ..., x_{i+(m-1)t}]^T, i = 1,2,...,M\} \qquad (1)$$

Where *m* is the numbers of points in Phase Space Reconstruction, $M=N-(m-1)t$, *m* and *t* are respectively delay time and embedding dimensions. Taken has proved that d is the dimensions when m>2d+1, so dynamical system of reconstruction is equivalent to original dynamical system.
At present, there are two kinds of method to choose the embedding dimensions *m* and time delay *t*. A sort of viewpoint think *m* and *t* are independence. The other viewpoint think *m* and *t* is correlate, time-window method or C-C method can be used to calculate them[5]. The paper using C-C method(Kim et al., 2009) to get embedding dimensions *m* and time delay *t*.

2.2 Identification of Time Series Chaos

The basic feature of chaotic motion is that the motion is very sensitive to initial conditions, it can measure the degree of separation. Two rail line on the close separates exponentially with time. Lyapunovi index is to quantitively describe the phenomenon based on whole. When the Lyapunov index is less than zero, the tracks shrink to point, when the Lyapunov index is greater than zero, the tracks separates, if trail has other factors, Chaotic Attractors will form based on the interaction. So the biggest Lyapunov index is positive is usually used as an important condition to judge chaos nature(Chen, 2005).

2.3 SimpleMKL

Kernel methods, such as support vector machines (SVM) have proved to be efficient tools for solving learning problems like classification or regression.Let $\{x_i, y_i\}_{i=1}^{\ell}$ be the learning set, where x_i is training specimen, y_i is objective value.
The solution of the learning problem is of the form:

$$f(x) = \sum_{i=1}^{\ell} a_i K(x, x_i) + b^* \qquad (2)$$

Where a_i and b^* are some coefficients to be learned from examples, while $K(\bullet,\bullet)$is a given positive definite kernel, $f(x)$ can be solved by a convex quadratic optimization.
The multiple kernels learning are Convex Combination of classical kernel in essence.

$$K(x, x') = \sum_{m=1}^{M} d_m K_m(x, x') \quad with \quad d_m \geq 0, \sum_{m=1}^{M} d_m = 1 \qquad (3)$$

Where *M* is the total number of kernels. k_m can simply be classical kernels (such as Gaussian kernels) with different parameters. d_m is weights of $K_m(x, x_i)$.

In the SVM methodology, the decision function is a form given in Formula(2), where the optimal parameters a_i and b^* can be obtained by solving the dual of the following optimization problem Formula(4):

$$\min_{f,b,\xi} \frac{1}{2}\|f\|_{\mathrm{H}}^2 + C\sum_i \xi_i$$

$$s.t.\ \ y_i\left(f\left(x_i\right)+b\right) \geq 1-\xi_i \quad \forall i$$

$$\xi_i \geq 0 \quad \forall i \tag{4}$$

In the MKL framework, one need look for a decision function of the form: $f(x)+b=\sum_m f_m(x)+b$,each function f_m associated with a kernel $K_m()$, Rakotomamonjy A. et al. (2008) propose to address the MKL SVM problem by solving the following convex problem based on (Formula 4):

$$\min_{\{f_m\},b,\xi,d} \frac{1}{2}\sum_m \frac{1}{d_m}\|f\|_{\mathrm{H}_m}^2 + C\sum_i \xi_i$$

$$s.t.\ \ y_i\left(\sum_m f_m\left(x_i\right)+y_i b\right) \geq 1-\xi_i \quad \forall i$$

$$\xi_i \geq 0 \quad \forall i$$

$$\sum_m d_m = 1, \quad d_m \geq 0 \quad \forall m \tag{5}$$

Formula (5) is the primal MKL problem, where each d_m controls the squared norm of f_m in the objective function. The MKL problem Formula (5) can equivalent to:

$$\min_d J\left(d\right) \ \text{ such that}$$

$$\sum_{m=1}^{M} d_m = 1, \quad d_m \geq 0, \tag{6}$$

where $J\left(d\right) = \min_{\{f\},b,\xi} \frac{1}{2}\sum_m \frac{1}{d_m}\|f\|_{\mathrm{H}_m}^2 + C\sum_i \xi_i$

$$s.t.\ \ y_i\left(\sum_m f_m\left(x_i\right)+y_i b\right) \geq 1-\xi_i$$

$$\xi_i \geq 0 \quad \forall i \tag{7}$$

The problem Formula(6) can be solved by a simple gradient method[34].The Lagrange of Formula (7) follows:

$$L = \frac{1}{2}\sum_m \frac{1}{d_m}\|f\|_{\mathrm{H}_m}^2 + C\sum_i \xi_i + \sum_I \alpha_i\left(1-\xi_i - y_i\sum_m f_m\left(x_i\right) - y_i b\right) - \sum_i v_i\xi_i + \lambda(\sum_m d_m - 1) - \sum_m \eta_m d_m \tag{8}$$

Where α_i and v_i are the Lagrange multipliers, when setting to zero the gradient of the Lagrangian with respect to the primal variables. The dual problem of Formula (7) can be obtained

$$\max_\alpha -\frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j \sum_m d_m k_m\left(x_i, x_j\right) + \sum_i \alpha_i$$

$$with \ \ \sum_i \alpha_i y_i = 0$$

$$c \geq \alpha_i \geq 0 \quad \forall i \tag{9}$$

Which is identified to the standard SVM dual formulation, only the kernel function becomes a combination of kernel functions:

$$K(x_i,x_j)=\sum_m d_m K_m(x_i,x_j) \tag{10}$$

Function $J(d)$ is defined as the optimal objective value of problem Formula (7). Because of strong duality, $J(d)$ is also the objective value of the dual problem Formula (11):

$$J(d)=-\frac{1}{2}\sum_{i,j}\alpha_i^*\alpha_j^* y_i y_j \sum_m d_m k_m\left(x_i,x_j\right)+\sum_i \alpha_i^* \tag{11}$$

Where α* can maximizes Formula(9), the objective value *J*(*d*) can be obtained by any SVM algorithm. So SimpleMKL can thus take advantage of any progress in single kernel algorithms.
The algorithm in the previous section focuses on binary classification, MKL algorithm can be extended to other regression with only little changes, Formula (7) can be changed:

$$J(d)=\min_{f_m,b,\xi_t}\frac{1}{2}\sum_m\frac{1}{d_m}\|f_m\|_{H_m}^2+c\sum_i\left(\xi_i+\xi_i^*\right)$$

$$s.t.\ y_i-\sum_m f_m\left(x_i\right)-b\le\varepsilon+\xi_i$$

$$\sum_m f_m\left(x_i\right)+b-y_i\le\varepsilon+\xi_i^*\quad\forall i$$

$$\xi_i\ge 0,\xi_i^*\le 0\quad\forall i \tag{12}$$

The dual problem of Formula (12) follows:

$$J(d)=\max_{\alpha,\beta}\sum_i(\beta_i-\alpha_i)y_i-\varepsilon\sum_i(\beta_i+\alpha_i)-\frac{1}{2}\sum_{i,j}(\beta_i-\alpha_i)\left(\beta_j-\alpha_j\right)\sum_m d_m k_m\left(xi,x_j\right)$$

$$with\ \sum_i\left(\beta_i-\alpha_i\right)=0$$

$$0\le\alpha,\beta\le c,\forall i \tag{13}$$

Where $\{\alpha_{I,}\ _{\beta i}\}$ are Lagrange multipliers, *J*(*d*) is differentiable, the problem can be soled by.

2.4 Forecasting based on phase space reconstruction and SimpleMKL

Step1 The delay time t and the embedding dimension m can be got by C-C methods.
Step2 Phase Space Reconstruction
Step3 Reconstruction matrix as training parameters, the system of prediction is developed after simpleMKL regression.
Step4 The next predictive value can be got by putting the predictive value to (2) and calculating (3).

3. Prediction for water inrush in mine and results

3.1 Data and Analysis

The second Liuqiao mine is located in Midwest of Huaibei coal field. The study area is a closed and semiclosed grid hydrogeological unit. So it can be considered that the study area is a continuous evolution system, in addition, due to mine discharge is the synthesis results of some factors, such as water pressure, the rock pressure, impermeable layer thickness, faults, underground mining etc. The interacion among these factors is complicated in the course of specific mine discharge(Yang and Chen, 2009). Mine discharge data trend in Figure (1) during January 1990 and February 2005 of Liuqiao second mine.

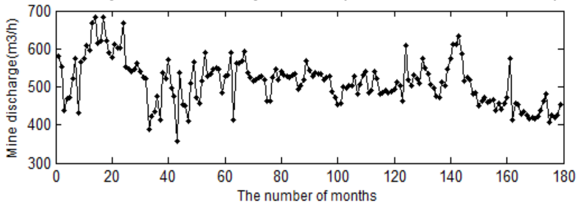

Figure 1: Mine discharge curve

3.2 Mine discharge forecasting and results analysis

The follow graph is got by using $C-C$ algorithm (Figure 2)

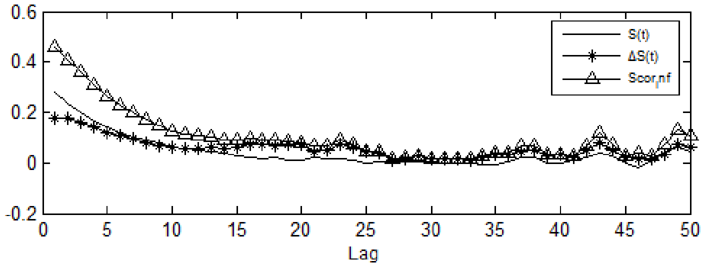

Figure 2: Time delay and embedding dimension based on C-C algorithm

Table 1: The accuracy analysis

Time	Actual values(m3/h)	Predictive values(m3/h)	Absolute error	Relative error
2004-8-1	462.3300	454.8056	7.5244	0.0163
2004-9-1	482.5800	467.5854	14.9946	0.0311
2004-10-1	406.8800	436.7100	29.8300	0.0733
2004-11-1	425.3100	457.2013	31.8913	0.0750
2004-12-1	420.4400	451.4993	31.0593	0.0739
2005-1-1	424.9400	447.7985	22.8585	0.0538
2005-2-1	454.7900	435.3154	19.4746	0.0428

The first minimal point of delta *△s(t)*: *t*=11, so optimal delay time *t*=11, the local minimum point of *Scor*: *t* =33, that is *tw*=33, according to *tw*=(*m*-1)·*tau* the embedding dimension *m*=4 can be got; maximum Lyapunov exponent λ=0.1844>0 by using wolf algorithm. So historical data of mine discharge has chaotic characteristics. So mine discharge can be prediction based Phase Spase Reconstruction and SimpleMKL.

A total of 179 data during January 1990 and February 2005, among them, Data during January 1990 and July 2004 as a training sample and get the prediction model, then to predict mine discharge during August 2004 and February 2005 and compared with the corresponding actual value, the absolute error and relative error as comparative indicators(Table 2).

In so far as some specific problems, there is not a consensus to choose better kernel parameters and weighted coefficient, in order to require optimal parameters, using usually empirical method, experimental contrast method, a large range search, cross validation method and so on.

In this test, parameters choosing as follow:

regularization parameters C is got using cross validation method:

regularization parameters: C=700;

kernel parameter: σ=[0.01 0.06 0.11 0.16 0.5 1 2 5 7 10 12 15 20]; *d*=[1,2,3];

The algorithm of reduced gradient method is terminated when a stopping criterion is met *DualGap*=0.01.

Table 1 show machine learing methods to predict mine discharge has a high precision, relative error control within 0.1, which further proves that the method based on Phase Space Reconstruction and SimpleMKL to forecast chaotic time series is feasible, processes a high precision and can be applied to practice.

4. Conclusion

This paper puts forward a kind of prediction method of mine discharge based on Phase Space Reconstruction and SimpleMKL.

The prediction method based Phase Space Reconstruction and simpleMKL is feasible and the precision is well, which can be applied to practical problem. Maximum lyapunov of mine discharge exponent λ=0.1844>0 in the field , so mine discharge has chaotic characteristics, thus input and output parameters can be obtained by the method of Phase Space Reconstruction and trained to get the prediction model, then mine discharge can be predicted.

By testing of Lorenz chaotic time series and prediction of mine discharge, it is proved that the method based on Phase Space Reconstruction and SimpleMKL to forecast mine discharge is feasible, prossesses good robustness, high precision and a strong practical value.

Acknowledgements

The work was supported by a grant awarded from National Natural Science Foundation of China (NSFC) grant (40472146).

Reference

Bach F.R., 2008, Consistency of the group Lasso and multiple kernel learning, Journal of Machine Learning Research, 9: 1179-1225.

Bach F.R., Lanckriet G.R.G., 2004, Multiple Kernel Learning, Conic Duality, and the SMO Algorithm, In: Proceedings of the 21st International Conference on Machine learning.

Chen Y.F., 2005, Application of chaotic time series in Hydrological Forecastingp[D], Chang`an University,

Cui W.Z., Zhu C.C., Bao W.X., Liu J.H.., 2005, Prediction of the chaotic time series usingsupport vector machines for fuzzy rule-based modeling, Acta Phys. Sin, 54(07): 3009-3018.

Chapelle O., Haffer P., Vapnik V.N., 1999, Support vector machines for histogram-based image classification, IEEE Transactions on Neural Networks, 10(5): 1055-1064

Gu W., Zhai D.H., 1995, A neural network for fast prediction of chaotic the series, Journal of Fuan University(Natural Science), 34(3): 262-268.

Genton M.G., 2001, Classes of kernels for machine learing: a statistics perspective, Journal of Machine Learning Research, 2: 299-312.

Kingsbury N., Tay D.B.H., Palaniswami M., 2005, Multi-scale kernel methods for classification, In: Proceedings of IEEE Workshop on Machine Learning for Signal Processing, 43-48.

Kim H.S., Eykholt R., Salas J.D., 1999, Nonlinear dynamics delay times and embedding windows[J], Physica D, 127: 48-60.

Meng Q.F., Peng Y.H., 2007,Improved adding weight first order local prediction method for chaotic time series[J], Computer Engineering and Applications, 43(35):61-64.

Packard N.H., Crutchfietd J.P., Farmer J.D., 1980, Geometry from a time series. Phys. Rev. Lett, 45(9):712-716.

Rakotomamonjy A., Bach F.R., Canu S., Grandvalet Y., 2008, SimpleMKL, Journal of Machine Learning Research, 9: 2491-2521.

Tang Y.M., Xiao J, 2006, Condition and mechanism about chaos formation in the evolution process of groundwater system of minning area. Journal of China Coal Society, 31(1): 45-49.

Takens F., 1981, Determing strange attractors in turbence[J],Lecture notes in Math, 898: 361-381.

Vapnik V.N., 1999, The Nature of Statistical Learning Theory, New York: Springer-Verlag.

Yang Z., Guo J., Xu W. Nie X.F., Wang J., Lei J.J., 2006, Multi-scale Support Vector machine for regression estimation, ISNN'06, LNCS, 3971:1030-1037.

Yang Y.G., Chen Y.H., 2009, Chaotic Characteristics and Prediction for Water Inrush in Mine[J], Journal of China University of Geosciences, 34(2):258:262.

Zhang C.W., Fu B., Xu Y., 2004, Research on gushing water in coal-mine[J].,Coal Technology, (5).

Zhang J.S., Xiao X.C., 2000, Prediction of chaotic time series by using adaptive higher order nonlinear fourier infrared filter[J]. Acta Phys. Sin., 49(07):1221-1227.