ISSN : 0974 - 7435 Volume 10 Issue 11

2014

BioTechnology
An Indian Journal

FULL PAPER
BTAIJ, 10(11), 2014 [5510-5517]

A genetic- support vector regression algorithm for oil field development and production prediction

Guo Dongfeng
Xinxiang University, Xinxiang 453003, (CHINA)
E-mail : guodongfeng1963@163.com

ABSTRACT

Accurate prediction of oil production is very important to help the company make a reasonable plan and avoid blind investment and achieve sustainable development. The selection of the appropriate parameters of support vector regression algorithm is very important for the forecasting performance of support vector regression algorithm. This study employs genetic algorithm to select the appropriate parameter of support vector regression algorithm.Thus, this paper presents genetic- support vector regression algorithm for oil field development and production prediction. The comparison of the oil production forecasting error among genetic-support vector regression algorithm shows that the oil production forecasting error of genetic-support vector regression algorithm is small than support vector regression algorithm and BP neural network.

KEYWORDS

Oil field; Oil production prediction; Regression algorithm; Genetic algorithm.

INTRODUCTION

Accurate prediction of oil production is very important to help the company make a reasonable plan[1-3] and avoid blind investment and achieve sustainable development[4,5]. Support vector regression algorithm is a popular prediction model[6,7]. It can map the input data into a high dimension space by nonlinear mapping function and solve the regression problem in the high dimension space[8]. The selection of the appropriate parameters C, σ, ε of support vector regression algorithm is very important for the forecasting performance of support vector regression algorithm. This study employs genetic algorithm to select the appropriate parameter of support vector regression algorithm. Thus, this paper presents genetic- support vector regression algorithm for oil field development and production prediction.

Oil production of Jianghan oil region from 1983 to 2005 is used to study the effectiveness of genetic- support vector regression algorithm compared with other intelligent prediction methods. In order to show the superiority of genetic-support vector regression algorithm further, the comparison of the oil production forecasting error among genetic-support vector regression algorithm, support vector regression algorithm and BP neural network is employed. The comparison of the oil production forecasting error among genetic-support vector regression algorithm shows that the oil production forecasting error of genetic-support vector regression algorithm is small than support vector regression algorithm and BP neural network.

GENETIC- SUPPORT VECTOR REGRESSION ALGORITHM

Support vector regression algorithm can map the input data into a high dimension space by nonlinear mapping function and solve the regression problem in the high dimension space. The regression function of support vector regression algorithm can be given as follows:

$$\mathbf{f(x) = w \cdot \phi(x) + b} \tag{1}$$

where x_i represents the input vector, w represents the weight vector and b represents the bias. We can solve the regression problem by introducing two positive slack variables ξ and ξ^*:

$$\min \left[\frac{1}{2}\|\mathbf{w}\|^2 + \mathbf{C}\sum_{\mathbf{i=1}}^{\mathbf{n}} \left(\xi_{\mathbf{i}} + \xi_{\mathbf{i}}{}^{*} \right) \right] \tag{2}$$

subject to

$$\begin{cases} y_i - \langle w \cdot \phi(x) + b \rangle \le \varepsilon + \xi_i \\ \langle w \cdot \phi(x) + b \rangle - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$

where C is the regularization parameter, ξ and ξ^* are the two positive slack variables.

Finally, the regression function of support vector regression model is obtained by solving the above optimization problem.

$$\mathbf{f(x)=\sum_{i=1}^{n}(a_i-a_i^*)k(x_i,x)+b} \qquad (3)$$

where a_i, a_i^* are the Lagrangian multipliers, and $k(x_i,x_j)$ is the kernel function.

The Gaussian kernel:

$$\mathbf{k(x_i,x)=exp\left(\frac{\|x_i-x\|}{2\sigma^2}\right)} \qquad (4)$$

is employed to establish the support vector regression model.

And support vector regression model with the Gaussian kernel is obtained by solving the above optimization problem.

$$\mathbf{f(x)=\sum_{i=1}^{n}\left(a_i-a_i^*\right)exp\left(\frac{\|x_i-x\|}{2\sigma^2}\right)+b} \qquad (5)$$

where σ is the width of the Gaussian kernel.

The selection of the appropriate parameters C,σ,ε of support vector regression algorithm is very important for the forecasting performance of support vector regression algorithm. This study employs genetic algorithm to select the appropriate parameter of support vector regression algorithm.

Genetic algorithm (GA), based on Darwinian evolution principle[9],has the higher global optimal ability[10]. Therefore, the detailed steps of selecting the appropriate parameters C,σ,ε of support vector regression algorithm can be described as follows:

Step 1 Code the parameters C,σ,ε of support vector regression algorithm

Step 2 Randomly generate a population of the chromosomes

Step 3 Set the training parameters of genetic algorithm

The probability of crossover is set to 0.6,and the probability of mutation is set to 0.01.

Step 4 Evaluate the fitness of each chromosome by the following formula:

$$\mathbf{F=\left[1+\frac{1}{L}\sum_{i=1}^{L}\left|\frac{y_i-y_i'}{y_i}\right|\right]^{-1}} \qquad (6)$$

where y_i is the actual value, and y_i' is the validation value.

Step 5 Generate the new chromosome by genetic operators

Step 6 The computational process terminates when the maximum iterative number is reached. Otherwise, go to Step 4.

EXPERIMENTAL RESULTS

As shown in Figure 1, oil production of Jianghan oil region from 1983 to 2005 is used to study the effectiveness of genetic- support vector regression algorithm compared with other intelligent prediction methods. Oil production data of Jianghan oil region from 1983 to 1998 are used to establish the training sample sets with 4 dimensions, and others are used to establish the testing sample sets. The training sample sets are described as follows:

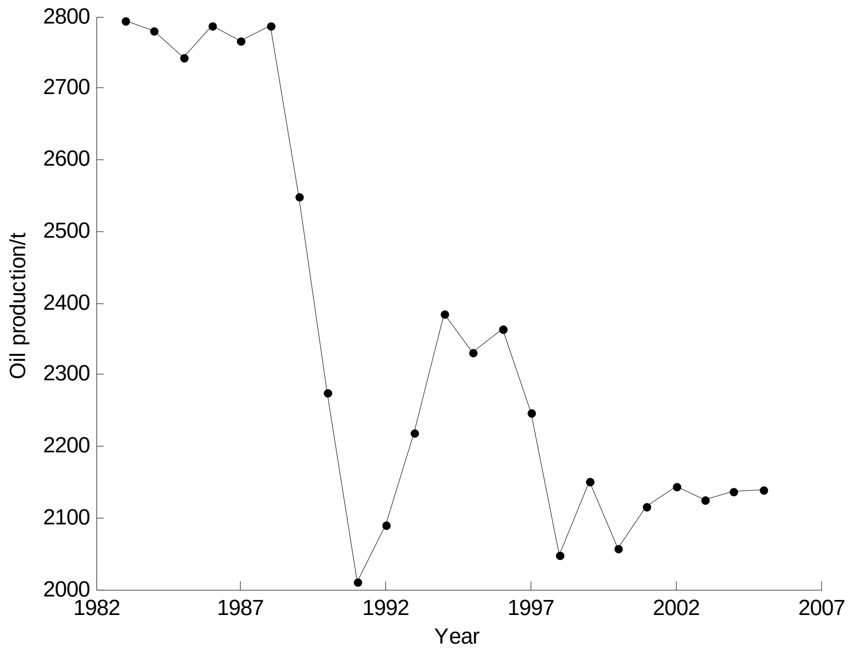

Figure 1 : Oil production of Jianghan oil region from 1983 to 2005

$$X=\begin{bmatrix} \mathbf{a}_{1983} & \mathbf{a}_{1984} & \mathbf{a}_{1985} & \mathbf{a}_{1986} \\ \mathbf{a}_{1984} & \mathbf{a}_{1985} & \mathbf{a}_{1986} & \mathbf{a}_{1987} \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_{1994} & \mathbf{a}_{1995} & \mathbf{a}_{1996} & \mathbf{a}_{1997} \end{bmatrix}, Y=\begin{bmatrix} \mathbf{a}_{1987} \\ \mathbf{a}_{1988} \\ \vdots \\ \mathbf{a}_{1998} \end{bmatrix} \quad (7)$$

where X represents the input vectors of the training sample sets and Y represents the corresponding outputs of the training sample sets.

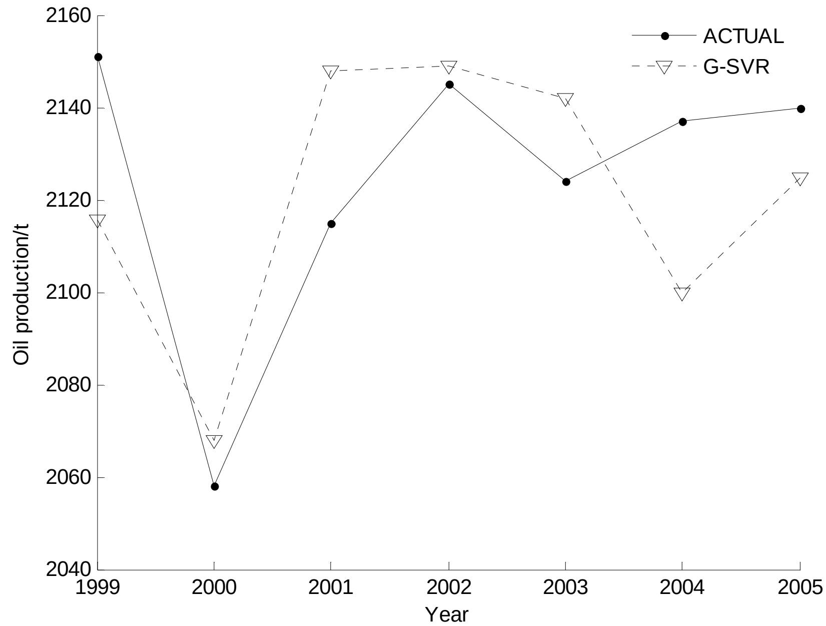

Figure 2 : The oil production forecasting value of genetic-support vector regression algorithm

Figure 2 shows the oil production forecasting value of genetic-support vector regression algorithm, the oil production forecasting value of support vector regression algorithm is given in Figure 3, and Figure 4 shows the oil production forecasting value of BP neural network.

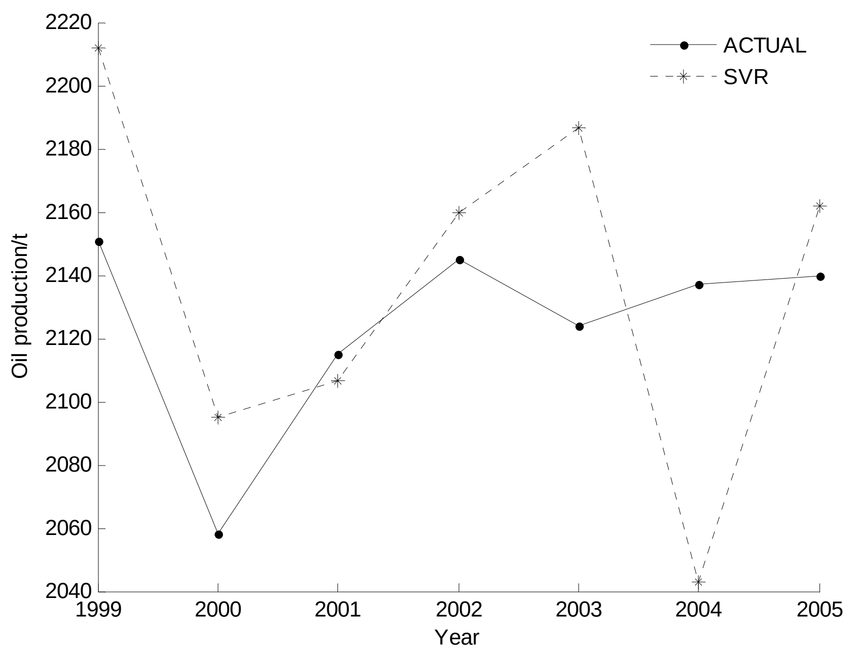

Figure 3 : The oil production forecasting value of support vector regression algorithm

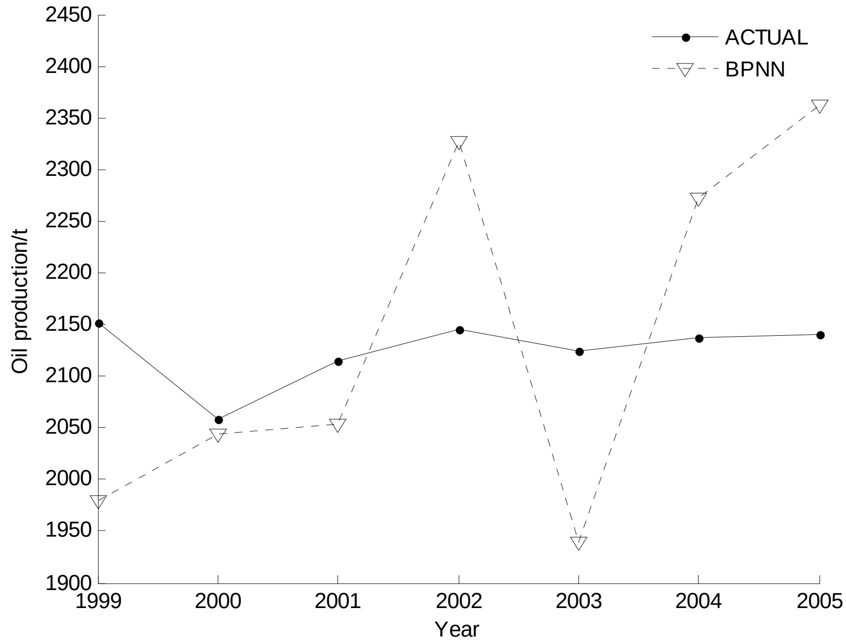

Figure 4 : The oil production forecasting value of BP neural network

In order to show the superiority of genetic-support vector regression algorithm further, the comparison of the oil production forecasting error among genetic-support vector regression algorithm, support vector regression algorithm and BP neural network is employed. Figure 5 shows the oil production forecasting error of genetic-support vector regression algorithm, the oil production forecasting error of support vector regression algorithm is given in Figure 6,and Figure 7 shows the oil production forecasting error of BP neural network.

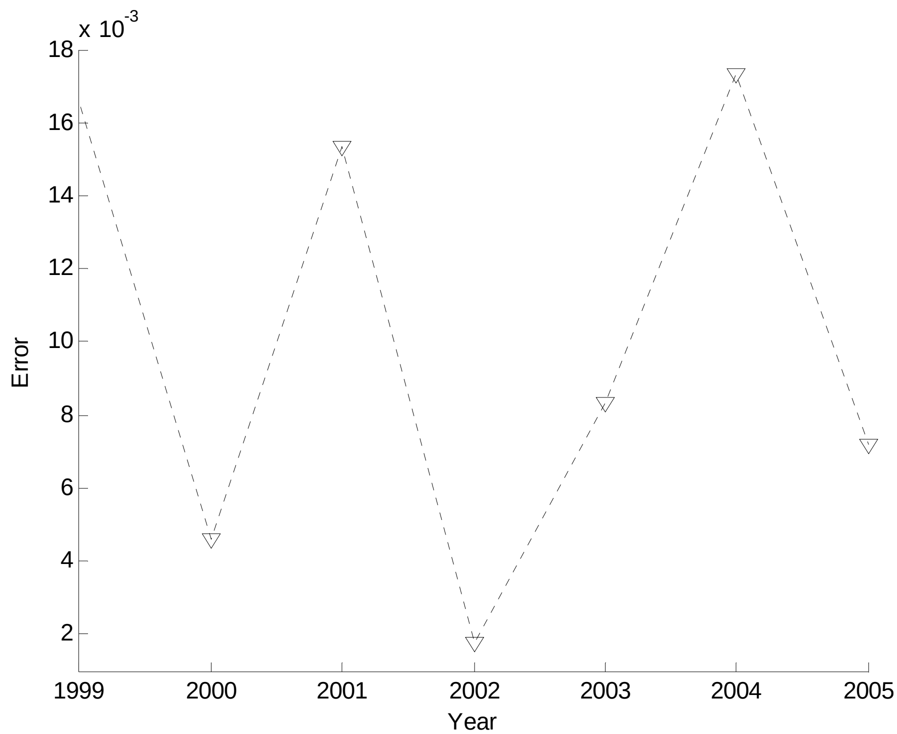

Figure 5 : The oil production forecasting error of genetic-support vector regression algorithm

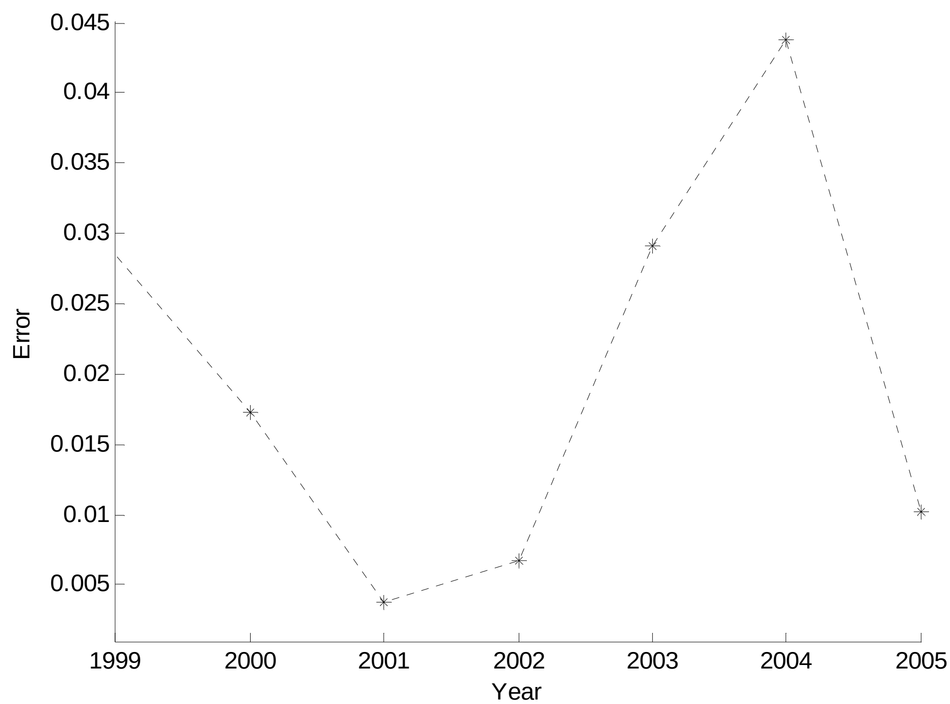

Figure 6 : The oil production forecasting error of support vector regression algorithm

The comparison of the oil production forecasting error among genetic-support vector regression algorithm shows that the oil production forecasting error of genetic-support vector regression algorithm is small than support vector regression algorithm and BP neural network.

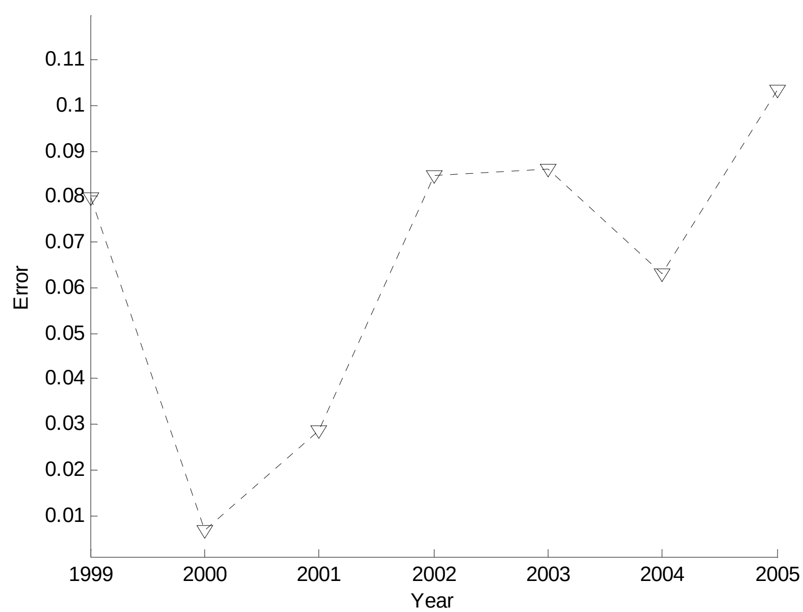

Figure 7 : The oil production forecasting error of BP neural network

CONCLUSION

This paper presents genetic- support vector regression algorithm for oil field development and production prediction, and this study employs genetic algorithm to select the appropriate parameter of support vector regression algorithm. In order to show the superiority of genetic-support vector regression algorithm further, he comparison of the oil production forecasting error among genetic-support vector regression algorithm, support vector regression algorithm and BP neural network is employed. The comparison of the oil production forecasting error among genetic-support vector regression algorithm shows that the oil production forecasting error of genetic-support vector regression algorithm is small than support vector regression algorithm and BP neural network.

REFERENCES

[1] Yu Deliang, Zhang Yongming, Bian Hongmei, Wang Xinmin, Qi Weigui; Producing Fluid Level Forecasting of A New Submersible Reciprocating Pump in Onshore Oil Wells, AISS, **4(7)**, 335-343 **(2012)**.

[2] Xiaodong Wu, Ming Zhu, Guoqing Han, Yongsheng An, Kai Liu; Wellbore Temperature Profile Prediction in Ultra-Heavy Oil Production Assisted by Light Oil in Ultra-Deep Wells, AISS, **4(10)**, 149-157 **(2012)**.

[3] Mikael Höök, Kjell Aleklett; A decline rate study of Norwegian oil production, Energy Policy, **36(11)**, 4262-4271 **(2008)**.

[4] X.Liu, G.Zhao, Y.C.Jin; Coupled reservoir/wormholes model for cold heavy oil production wells, Journal of Petroleum Science and Engineering, **50(3–4)**, 258-268 **(2006)**.

[5] Jianliang Wang, Lianyong Feng, Lin Zhao, Simon Snowden, Xu Wang; A comparison of two typical multicyclic models used to forecast the world's conventional oil production, Energy Policy, **39(12)**, 7616-7621 **(2011)**.

[6] Somnath Nandi, Yogesh Badhe, Jayaram Lonari, U.Sridevi, B.S.Rao, Sanjeev S.Tambe, Bhaskar D.Kulkarni; Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: study of benzene isopropylation on Hbeta catalyst, Chemical Engineering Journal, **97(2–3)**, 115-129 **(2004)**.

[7] S.G.Patil, S.Mandal, A.V.Hegde; Genetic algorithm based support vector machine regression in predicting wave transmission of horizontally interlaced multi-layer moored floating pipe breakwater, Advances in Engineering Software, **45(1)**, 203-212 **(2012)**.

[8] Nikolaos Kavvadias, Vasiliki Giannakopoulou, Spiridon Nikolaidis; Development of a customized processor architecture for accelerating genetic algorithms, Microprocessors and Microsystems, **31(5)**, 347-359 **(2007)**.

[9] Dhiranuch Bunnag, Min Sun; Genetic algorithm for constrained global optimization in continuous variables,Applied Mathematics and Computation, **171(1)**, 604-636 **(2005)**.