Predicting the oil field production using the GMC(1, n) model

Xuemei Hu [a], Bijue Jia [a]

School of Science, Southwest Petroleum University, Chengdu 610500, China

[a]787565373 @qq.com

Abstract

The production forecast of the oil field is of importance in the petroleum engineering, which can help the engineers to make the working schedules, adjust the future plans and make the investment decisions, etc. In this paper we introduce a novel grey prediction model, the GMC(1, n) model, to predict the oil field production, which can be built with small samples. The case study has been carried out with the data from a real world oil field in north China, and the GMC(1, n) has been shown to outperform the traditional GM(1, n) model with high precision in production forecast, which implies the high potential of the novel model to predict the oil field production.

Keywords

Oil field production, Grey System Theory, Grey prediction model, Small samples.

1. Introduction

Estimating the future performance of the petroleum reservoir is an important task in the petroleum engineering. In the previous studies, the decline curve methods (DCM) [1] and artificial intelligent methods (AIM) [2, 3] were often used for the prediction of the petroleum production. However, a large scale of historical data is often needed in the previous methods, which will bring large amount of workload for the petroleum engineers and even sometimes the forecast precision is not applicable. In this study, we will introduce a novel prediction model based on the grey system theory, which can be built on small samples and is also accurate.

The grey system theory(GST) was pioneered by Deng[4], which was developed from idea of "Grey Box". The GST contains a family of tools which are efficient to solve the problems with small samples. The grey prediction models are of high importance to the GST, which have been applied in various fields [5, 6]. The GM(1, n) model is one of the most important prediction models in the grey prediction models as it can been deemed as a general form of the other grey prediction models, such as the so called GM(1, 1) [7], the GM(1, 1, k) model[8], etc. However, the traditional GM(1, n) model has been pointed out to be a wrong model in the research of Tien [9] recently, and the a novel model called GMC(1, n) model has been proposed. The GMC(1, n) can be regarded as an improved model of the GM(1, n), and it has been shown to perform much better than the GM(1, n) in the indirect measurement in the research of Tien. Being similar to the GM(1, n), the GMC(1, n) is one kind of multiple regression models, and it can be built with small samples.

In this study, the GMC(1, n) will be used to predict the oil field production. In the rest of this paper, we will firstly overview the principles of the GMC(1, n) briefly, and the case study will be carried out with the production data from a real world oil field in north China, in which the performances of the GMC(1, n) along with the GM(1, n) will be compared and analyzed. The conclusions and perspectives will be drawn in the last part of this paper.

2. The foudamental principles of GMC(1, n) model

2.1 The whitening equation of the GMC(1, n) model

Set the original sequence as $\{x^{(0)}(k) \mid k=1,2,...,n\}$, the first order accumulative generation operation (1-AGO) of the original sequence is defined as $\{x^{(1)}(k) \mid x^{(1)}(k)=\sum_{m=1}^{k} x^{(0)}(m),\ k=1,2,...,n\}$.

The first order differential equation

$$\frac{dx_1^{(1)}(t)}{dt}=ax_1^{(1)}(t)+f(t) \tag{1}$$

is called the whiten equation of the GMC(1, n) model, where $f(t)=\sum_{i=2}^{N} b_i x_i^{(1)}(t)+u$. In the Eq.(1), the sequence $x_1^{(1)}(t)$ is often called the feature sequence, and the other sequences are called the reliance sequences.

2.2 Parameters estimation of the GMC(1, n) model

The parameters in Eq.(1) can be obtained by solving the following linear equation

$$(B^T B)^{-1}\alpha=B^T Y, \tag{2}$$

where

$$\alpha=[a,b_2,b_3,...,b_N,u]^T,$$

$$B=\begin{bmatrix} -\frac{1}{2}(x_1^{(1)}(1)+x_1^{(1)}(2)), & \frac{1}{2}(x_2^{(1)}(1)+x_2^{(1)}(2)), & ..., \frac{1}{2}(x_2^{(1)}(1)+x_2^{(1)}(2)), & 1 \\ -\frac{1}{2}(x_1^{(1)}(2)+x_1^{(1)}(3)), & \frac{1}{2}(x_2^{(1)}(2)+x_2^{(1)}(3)), & ..., \frac{1}{2}(x_2^{(1)}(2)+x_2^{(1)}(3)), & 1 \\ ... & & & \\ -\frac{1}{2}(x_1^{(1)}(n-1)+x_1^{(1)}(n)), & \frac{1}{2}(x_2^{(1)}(n-1)+x_2^{(1)}(n)), & ..., \frac{1}{2}(x_2^{(1)}(n-1)+x_2^{(1)}(n)), & 1 \end{bmatrix}$$

and

$$Y=\left[x_1^{(0)}(2),x_1^{(0)}(3),...,x_1^{(0)}(n)\right]^T.$$

2.3 Response function and stored value

The continuous form of the response function of the GMC (1, n) model is given as follows by solving the whiten equation (1):

$$\hat{x}_1^{(1)}(t)=x_1^{(1)}(1)e^{-a(t-1)}+\int_1^t e^{-a(t-\tau)}f(\tau)d\tau. \tag{3}$$

The integration in Eq.(3) is in the interval $[1,t]$ because in the grey system theory it is often counted from 1. Discretizing the integrations in Eq.(3) using the trapezoid formula and noticing that $x^{(1)}(1)=x^{(0)}(1)$, the response function can be given as

$$\hat{x}_1^{(1)}(k)=x_1^{(0)}(1)e^{-a(t-1)}+\frac{1}{2}e^{-a(t-1)}f(1)+\sum_{\tau=2}^{k-1}[e^{-a(t-1)}f(\tau)]+\frac{1}{2}f(k). \tag{4}$$

By applying the first order inverse accumulative generation operation (1-IAGO), we have the stored value

$$\hat{x}_1^{(0)}(k)=\hat{x}_1^{(1)}(k)-\hat{x}_1^{(1)}(k-1), \tag{5}$$

where $k \geq 2$, and $\hat{x}_1^{(0)}(1) = x_1^{(0)}(1)$ when $k = 1$. The Eq.(4) and (5) are used to simulate and predict the feature sequence.

3. Case Study

3.1 Raw data collection

The raw data is collected from a real world oil field in North China, which is called the RQ oil field. The production data from the year of 2000 to 2012 has been collected for this case study, which is shown in Table 1. The oil production ($10^4 m^3$) is regarded as the feature sequence, while the number of oil wells, water wells, operations and the amount of water injections ($10^4 m^3$) are regarded as the reliance sequences for modelling, respectively. The data from 2000 to 2007 is used to build the prediction models, and that from 2008 and 2012 is used to test the prediction accuracy. Additionally, the traditional GM (1, n) model is also applied in this case study in order to compare the precision of the prediction models.

3.2 Evaluation criteria for modelling accuracy

The mean absolute percentage error (MAPE) is used to evaluate the modelling accuracy in this case study, which is defined as

$$MAPE = \frac{1}{p}\sum_{k=1}^{p} \left| \frac{\hat{x}_1^{(0)}(k) - x_1^{(0)}(k)}{x_1^{(0)}(k)} \right| \times 100\%, \tag{6}$$

where p stands for the total number of the simulation or prediction step.

Table 1. The production data of RQ oil field from the year of 2000 to 2012

Year	Oil wells	Water wells	Injections	Operations	Oil production
2000	2921	876	2190.534	516	456.1391
2001	3159	970	2217.963	494	450.7212
2002	3157	1012	2105.045	430	438.0026
2003	3283	1061	2102.293	490	435.2036
2004	3330	1128	2176.658	529	432.2888
2005	3404	1213	2198.899	552	435.1019
2006	3602	1211	2497.85	537	440.1115
2007	3803	1330	2744.744	489	447.0053
2008	3938	1397	3069.98	530	441.7077
2009	3926	1409	3197.818	570	425.0274
2010	3785	1490	3347.643	599	424.6889
2011	3483	1425	3446.377	699	419.8934
2012	3593	1620	3221.121	736	417.8266

3.3 Experiment results

The experiment result is shown in Table 2. The MAPE for simulation of GM (1, n) and GMC(1, n) is 19.84% and 0.09%, respectively. And the MAPE for prediction of GM (1, n) and GMC(1, n) is 174.15% and 9.96%, respectively. The maximum error of GM (1, n) is 323.78%, while that of GMC(1, n) is 12.43%.

It is obviously that the performance of GMC (1, n) is much better than the GM(1, n), and the performance of GM(1, n) cannot be applicable at all. This is because of the wrongness of the modelling procedures in the GM (1, n), which has been revealed by Tien[9]. The performance of the GMC (1, n) is applicable, which implies the applicability of the GMC(1, n) in predicting the oil field production.

Table 2. Experiment results of GMC (1, n) and GM(1, n) model

Year	Production	GMC(1,n)	Error(%)	GM(1,n)	Error(%)
2000	456.14	456.14	0.00	456.14	0.00
2001	450.72	450.73	0.00	447.07	0.81
2002	438.00	438.51	0.12	516.01	17.81
2003	435.20	435.12	0.02	557.40	28.08
2004	432.29	432.65	0.08	592.26	37.01
2005	435.10	436.30	0.27	488.58	12.29
2006	440.11	439.95	0.04	674.62	53.28
2007	447.01	447.83	0.18	489.10	9.42
2008	441.71	460.87	4.34	759.45	71.93
2009	425.03	467.25	9.93	993.24	133.69
2010	424.69	472.98	11.37	1118.90	163.46
2011	419.89	469.18	11.74	1779.42	323.78
2012	417.83	469.76	12.43	1161.01	177.87

4. Conclusions and perspectives

The GMC (1, n) model has been applied to predict the oil field production. The case study has been carried out with a real world oil field in China. The experiment results indicated that the GMC (1, n) model can be built with a few data, and the it

outperformed the traditional GM(1, n). Thus it can be summarized that the GMC (1, n) is of high potential to predict the oil field production accurately.

However, the prediction accuracy might be improved with more influence factors. In this study, only four influence factors have been chosen for the reliance sequences, and it can be seen that the prediction error has become larger with more prediction step. The future work will be oriented in selecting the influence factors based on the grey reliance theories.

Acknowledgements

This research was supported by the natural fund of education department of Sichuan Province (No.14ZB0388), the China Postdoctoral Science Foundation funded project, No: 2014M562509XB, and the Scientific Project of Sichuan Provincial Education Department (No. 15ZB0447)..

References

[1] J.J. Arps, Analysis of decline curves, Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers. Vol. 160 (1945), p. 228–247.

[2] Y. H. Zhong, L. Zhao, Z. Liu, Y. Xu and R. Li, Using a support vector machine method to predict the development indices of very high water cut oilfields, Petroleum Science, vol. 7 (2010), p. 379–384.

[3] S. Gupta, F. Fuehrer and B.C. Jeyachandra, *et al*, Production forecasting in unconventional resources using data mining and time series analysis, SPE/CSUR Unconventional Resources Conference–Canada, Society of Petroleum Engineers, 2014.

[4] J.L. Deng, Control problems of grey systems, System Control Letters. vol. 1 (1982), p. 288–294.

[5] M. Guo, J. Lan, Z. Lin and X. Sun, Traffic Flow Data Recovery Algorithm Based on Gray Residual GM(1, N) Model, J Transpn Sys Eng & IT. vol.12 (2012), p. 42-47.

[6] T.L. Tien, The indirect measurement of tensile strength of material by the grey prediction model GMC(1, n), Meas. Sci. Technol. vol.16 (2005), p. 1322-1328.

[7] S.F. Liu and Y. Lin, Grey Systems: Theory and Applications (Springer, Germany, 2010).

[8] J. Cui, S.F. Liu, B. Zeng and N.M. Xie, A novel grey forecasting model and its optimization, Applied Mathematical Modelling. vol.37 (2013), p. 4399–4406
[9] T.L. Tien, A research on the grey prediction model GM (1, n), Applied Mathematics and Computation. vol.218 (2012), p. 4903–4916.