American Journal of Computer Science and Technology
2024; Vol. 7, No. 2, pp. 38-42
https://doi.org/10.11648/j.ajcst.20240702.12

Research Article

A Method for Solving the Generalized Weng Model

Zeng Hongwu[1], Xu Qiuyu[2, *]

[1]School of Medical Informatics, Chongqing Medical University, Chongqing, China
[2]Experimental Teaching Center, Chongqing Medical University, Chongqing, China

Abstract

The Generalized Weng Model is one of the basic models for oil production forecasting. Professor Chen Yuanqian first proposed the linear iterative trial-and-error method to solve the generalized Weng Model, and scholar Zhao Lin proposed the method to solve the Weng model based on binary regression. In this paper, a new method for solving Weng Model is put forward. Taking Liaohe Oilfield in China as an example, the process and results of the three methods are compared, and the advantages and disadvantages of the three methods are analyzed. The results show that when the original linear iterative trial and error method solves the model, it needs to simulate the value of parameter b with computer software, and then select a judgment criterion to find the optimal b value. In this paper, a method based on binary regression is proposed which can directly calculate parameter *b*. The new method can directly calculate the parameter *b* better than the method based on binary regression. The method in this paper is to fit all the data at one time, avoiding the above two kinds of uncertainties, and the calculation workload is small and can be realized by EXCEL, which is convenient for technical personnel.

Keywords

Parameter, Prediction, Generalized Weng Model

1. Introduction

The famous geophysicist, Mr. Weng, proposed the Poisson Cycle model for predicting the reserves and medium- and long-term production of oil and gas fields in his monograph "Fundamentals of Prediction Theory" which is published in 1984. [1] In his monograph "Theory of Forecasting" published in 1991, Mr. Weng renamed the Poisson cycle model as the Life Cycle model, which was called the Weng model by later generations. [2] In 1987, Professor Zhao Xudong put more than 150 domestic and foreign oil and gas fields into the model for trial calculation, and the final result showed that it was feasible to apply the Weng's prediction model in the production and final recoverable reserves of oil and gas fields. [3-6]

In 1996, Professor Chen Yuanqian explored the establishment and derivation process of Weng's model, obtained the generalized Weng's model with a wider range of application, and proposed the linear test method to solve the parameters of the generalized Weng's model. Through verification, this method can predict the cumulative production, production, recoverable reserves, and maximum annual production of oil and gas fields. [7-14] In 2009, Lin Zhao presented a method for solving the generalized Weng model based on binary regression. [15]

In this paper, we will propose an estimation method for the

*Corresponding author: 101852@cqmu.edu.cn (Xu Qiuyu)

Received: 2 April 2024; **Accepted:** 22 April 2024; **Published:** 17 May 2024

parameters of the generalized Weng model based on nonlinear regression, which can directly obtain the parameter b that is better than the solution based on the binary regression method, and can improve the fitting and prediction accuracy of the generalized Weng model.

2. Estimation Method for Solving the Parameters of the Generalized Weng Model Based on Nonlinear Regression

In the generalized Weng model, it is necessary to find the hidden regularity from a large amount of data, and the generalized Weng model belongs to a nonlinear model. Compared with the linear model, the complexity of the nonlinear model increases geometrically, so the regression parameters of the linear model are easier to calculate, while the regression parameters of the nonlinear model are not easy to calculate. When solving practical problems, we often have to settle for the next best thing and use the approximate regression model of the nonlinear model to obtain the regression parameters of the nonlinear model. There are many ways to use the approximate regression model of nonlinear model to obtain the regression parameters of the nonlinear model. The method used in practical textbooks is to perform variable substitution on the nonlinear model, convert it into a linear model, and then use the linear regression model to obtain the regression parameters. This approach may add interference information, or lose the original information, and in the most extreme cases, the prediction accuracy of the derived nonlinear regression model will be greatly reduced. What is about to be discussed below is a simple way that can overcome these drawbacks.

Without losing generality, the form of the nonlinear function model can be expressed as follows

$$\left.\begin{aligned} y = f(x_1, x_2, \cdots, x_m, \alpha_1, \alpha_2, \cdots, \alpha_l) + e \\ e \sim N(0, \sigma^2) \end{aligned}\right\} \tag{1}$$

About the independent variables $x = (x_1, x_2, \cdots, x_m) \in R^m$ in the nonlinear function model (1), obviously, it is a point in the m-dimensional space, regarding the parameters $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_l) \in R^l$ in the nonlinear function model (1), apparently it is a point in the l-dimensional space, with regard to the dependent variable $y \in R^1$ in the nonlinear function model (1), it is obviously a point in one-dimensional space. The nonlinear function model (1) $f(x_1, x_2, \cdots, x_m; \alpha_1, \alpha_2, \cdots, \alpha_l)$ is also a multivariate function containing the parameters α, which is mostly a nonlinear function under normal circumstances. The purpose of nonlinear regression analysis is to obtain the best estimate of the parameter α in the l-dimensional space.

Suppose it is necessary to find the optimal estimate of the parameter α ($\hat{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_l) \in R^l$) of the nonlinear function in the l-dimensional space from a set of original experimental data $\{((x_1, x_2, \cdots, x_m)_i, y_i) \mid i = 1, 2, \cdots, n\}$:

$$S(\hat{\alpha}) = \min_{\alpha} S(\alpha) = \min_{\alpha} S(\alpha_1, \alpha_2, \cdots, \alpha_l)$$

during the above ceremony

$$S(\alpha) = \sum_{i=1}^{n} (y_i - f(x_1, x_2, \cdots, x_m; \alpha_1, \alpha_2, \cdots, \alpha_l))^2$$

Suppose that the set of the original experimental data of the nonlinear function model (1) in the m+1-dimensional space is:

$$\{((x_1, x_2, \cdots, x_m)_i, y_i) \mid i = 1, 2, \cdots, n\}$$

According to the original experimental data, the desired nonlinear function model is obviously located in the m+1-dimensional space X-Y. Why don’t we suppose that the predicted values of the nonlinear function model are as follows:

$$\{((x_1, x_2, \cdots, x_m)_i, \hat{y}_i) \mid i = 1, 2, \cdots, n\}$$

In order to directly find the predicted value of this nonlinear function model in the m+1-dimensional space X-Y, it is necessary to use large-scale software. Practical textbooks usually use some simple alternative methods that change the nonlinear function model into a linear function model for linear regression, find the parameter α of the nonlinear function, and then convert the nonlinear function model.

In this conversion process, the data set in X-Y space must first be transformed into a data set in X-Z space. Practical textbooks usually substitute variables as follows:

$$z = F(y) \tag{2}$$

The above variable substitution transforms the data set in X-Y space into a data set in X-Z space, that is:

$$\{((x_1, x_2, \cdots, x_m)_i, z_i) \mid i = 1, 2, \cdots, n\} = \{((x_1, x_2, \cdots, x_m)_i, F(y_i)) \mid i = 1, 2, \cdots, n\}$$

The data set corresponding to the set of predicted values in the data set in the X-Z space is the data set on the m+1-dimensional new variable space X-Z as:

$$\{((x_1, x_2, \cdots, x_m)_i, \hat{z}_i) \mid i = 1, 2, \cdots, n\} = \{((x_1, x_2, \cdots, x_m)_i, F(\hat{y}_i)) \mid i = 1, 2, \cdots, n\}$$

Some textbooks or papers use the following formula to find the sum of squares of the residuals of the data that set in X-Z space, namely:

$$S_1 = \sum_{i=1}^{n} (z_i - \hat{z}_i)^2 \tag{3}$$

On the basis of the residual sum of squares formula, the best estimate $\hat{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_l) \in R^l$ of the parameters α *is* obtained by using the least squares method, and then substituted α into equation (1) to obtain a nonlinear function model.

After a large number of experimental verifications, it was found that although a nonlinear function model can be obtained by using this method, this nonlinear function model quite naturally hides a serious defect that is not easy to detect. The sum S_2 of squares of the residuals of the data in X-Y space for the original m+1-dimensional variable is expressed by the following equation:

$$S_2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{4}$$

Although it can meet the minimum of the sum S_1 of the residuals of the square of the data in the X-Z space of the new m+1-dimensional variable, it may not be able to meet the minimum of the sum S_2 of the residuals of the square of the data in the X-Y space of the original m+1-dimensional variable, which will lead to the large error of the regression parameters of the obtained nonlinear mathematical model, and even lead to the complete failure of the obtained nonlinear mathematical model in serious cases.

For functions with unknown parameter α

$$\hat{z}_i = F(\hat{y}_i) = F[y(x_i)]$$

Expand by Taylor series at y_i

$$\begin{aligned} \hat{z}_i = F(\hat{y}_i) &= F(y_i) + F'(y_i)(\hat{y}_i - y_i) + \frac{1}{2} F''(y_i)(\hat{y}_i - y_i)^2 + \cdots \\ &= z_i + F'(y_i)(\hat{y}_i - y_i) + \frac{1}{2} F''(y_i)(\hat{y}_i - y_i)^2 + \cdots \end{aligned}$$

When the nonlinear mathematical model satisfies the condition that $\hat{y}_i \to y_i$, the term $(\hat{y}_i - y_i)$ in the above equation is an infinitesimal quantity when $\hat{y}_i \to y_i$. When satisfying $\hat{y}_i \to y_i$, we can omit all the higher-order infinitesimal terms of infinitesimal $(\hat{y}_i - y_i)$ and get:

$$(\hat{y}_i - y_i) \approx \frac{(\hat{z}_i - z_i)}{F'(y_i)} \tag{5}$$

Now we can organize the approximate expression of the sum S_2 of squares of the data residuals on the original m+1-dimensional variable space X-Y as:

$$S_2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \approx \sum_{i=1}^{n} \frac{(z_i - \hat{z}_i)^2}{[F'(y_i)]^2} \tag{6}$$

During the above ceremony: $z_i = F(y_i)$, $\hat{z}_i = F[y(x_i)]$, $F'(y_i) = F'(y)|_{y=y_i}$.

The least squares method is applied to equation (6) to find the approximate expression of the sum S_2 of squares of the data residuals on the original variable space X-Y, and then the best estimate of the parameter α in the l-dimensional space can be obtained, and the parameter α only need to be substituted into equation (1) to obtain the nonlinear mathematical model.

For the generalized Weng model that $Q=at^be^{-ct}$, the approximate regression method based on the nonlinear model is used to estimate the parameters of the generalized Weng model as follows. Make that:

$$P=F(Q)=\ln Q=\ln a+b\ln t-ct$$

In the above equation, the dependent variable P is derived from the independent variable Q, that is:

$$P'=1/Q$$

$$\begin{aligned} S_2 &= \sum_{i=1}^{n} (Q_i - \hat{Q}_i)^2 \approx \sum_{i=1}^{n} \frac{(P_i - \hat{P}_i)^2}{[F'(Q_i)]^2} \\ &= \sum_{i=1}^{n} Q_i^2 (\ln a + b \ln t_i - ct_i - \ln \hat{Q}_i)^2 \end{aligned}$$

Using the original data, we find the partial derivative of the above equation and solve the parameters a, *b*, *and c*, and then obtain the generalized Weng model corresponding to the practical problem.

3. A Case Analysis of Solving the Parameters of the Generalized Weng Model

We used the production data of Liaohe Oilfield from 1970 to 2006 cited in Zhao Lin's paper as an example for comparative analysis. Figure 1 is the curve trend diagram of the simulated data calculated by the statistical data and its three methods, and the data processing and graphics are done by EXCEL.

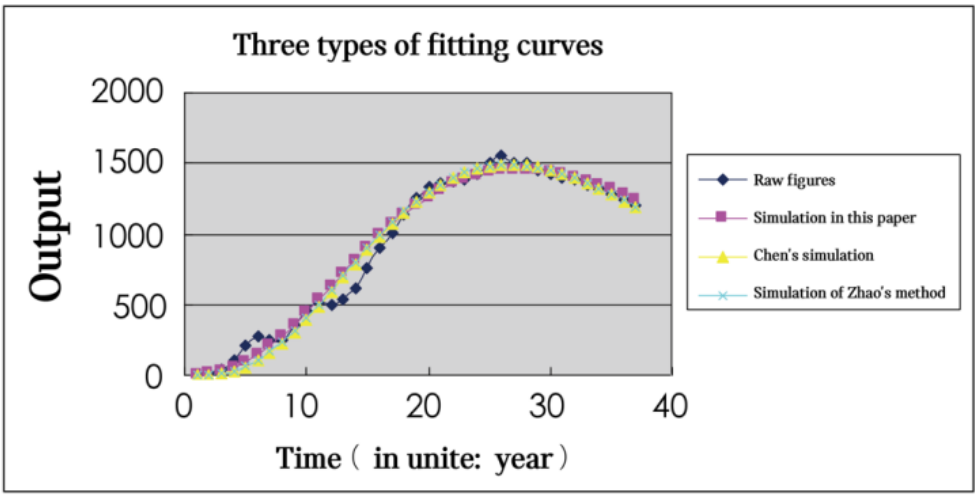

Figure 1. *Curve trend chart of statistical data and its simulated data.*

The following table shows the results of the parameters a, b and c, the maximum residuals squared, and the sum of squares of the residuals obtained by the three methods.

Table 1. *Parameters obtained by the three methods.*

Methods	Parameter *a*	Parameter *b*	Parameter 1/*c*	Maximum Residual Square	Sum of squares of residuals
Chan Method	0.2733	3.77	7.0664	31633.99021	180993.0375
Zhao Method	0.3554	3.654	7.2898	36708.91391	188722.0685
The Method of this article	0.980591406	3.156707527	8.706048648	41486.35276	202498.7392

Considering the maximum relative error and the sum of squares of the residuals, it can be seen that there is no statistical difference between the models obtained by the three methods.

4. Discussion

Professor Chen's solution to the generalized Weng model prediction is a base-value prediction, and when the linear iterative test method is used to solve the generalized Weng model, there will be multiple solutions because of the uncertainty in the value of parameter b and the selection of regression segments. When using the linear iterative test method to solve the Weng's model, it is necessary to use computer software to simulate the value of parameter b, and then select "the value b of the historical fitting of the highest correlation coefficient and the best yield and cumulative yield". Selecting the data segment with the best linear relationship in the uncertainty of the regression segment, and then the judgment process is more intuitive. Zhao Lin's method separates these two kinds of uncertainty when solving, and directly obtains parameter *b* through binary regression on the uncertainty of parameter b, but the uncertainty of regression segment selection is not as intuitive as Professor Chen's method, which it also relies on human judgment selection. The method in this article is a one-time fitting of the whole data, avoiding the two uncertainties mentioned above, and it has less computational work and can be implemented with EXCEL, which is convenient for technicians to use.

5. Conclusion

In this paper, a new method for solving Weng's model is proposed, and the advantages and disadvantages of the other three methods are analyzed and compared. At the same time, a new method based on binary regression is proposed. The new method can directly calculate the parameter b better than the method based on binary regression. This method is to fit all the data at one time, avoid the above two kinds of uncertainties, and the calculation workload is small and can be realized by EXCEL, which is convenient for technical personnel.

Abbreviations

EXCEL: Microsoft Office Excel

Author Contributions

Zeng Hongwu: Project administration, Resources
Xu Qiuyu: Conceptualization, Formal Analysis, Writing - original draft, Writing - review & editing

Conflicts of Interest

The article is entirely original, without any plagiarism or improper quotation; Submissions are made in the course of study, research or work without any conflict of interest.

References

[1] Weng Wenbo. *Fundamentals of Prediction Theory* [M], Beijing: Petroleum Industry Press, 2020. 67-86.

[2] Weng Wenbo. *Theory of Forecasting* [M]. International Academic Publishers: A Pergamon-CNPIEC Joint Venture, 1991.

[3] Zhao Xudong. *Methods for predicting oil and gas field production and final recoverable reserves* [J]. Petroleum Exploration and Development, 2019, 13(2): 72-78.

[4] Zhao Xudong. *Prediction of finite life systems using the Weng cycle model* [J]. Chinese Science Bulletin, 2018, 32(18): 1406-1409.

[5] Zhao Xudong. *Quantitative evaluation of petroleum resources* [M]. Beijing: Geological Publishing House, 2917.

[6] Zhao Xudong. *Introduction to Petroleum Mathematical Geology* [M]. Beijing: Petroleum Industry Press, 2006.

[7] Chen Yuanqian, Hu Jianguo, Zhang Dongjie. Logistic *Model derivation and autoregressive methods* [J]. Xinjiang Petroleum Geology, 2016, 17(2): 150-155.

[8] Chen Yuanqian. *Derivation and Application of Generalized Weng's Prediction Model* [J]. Natural Gas Industry, 2019, 16(2): 22-26.

[9] Chen Yuanqian, Hu Jianguo. *Review and new derivation of the original modeling of Weng's model* [J]. China Offshore Oil and Gas (Geology), 2020, 10(5): 317-324.

[10] Chen Yuanqian, Li Xuan. *Modern reservoir engineering* [M]. Beijing: Petroleum Industry Press, 2018.

[11] Chen Yuanqian. *Practical methods for oil and gas reservoir engineering* [M]. Beijing: Petroleum Industry Press, 2020, 1(6).

[12] Chen Yuanqian. *Derivation of the water drive curve relation* [J]. Acta PetroleiSinica, 2021, 6(2), 69-78.

[13] Chen Yuanqian. *Theoretical Derivation and Application of Nazarov's Empirical Formula for Determining Dead Reserves, Petroleum Exploration and Development* [J], 2020, 22(3), 63-68.

[14] Chen Yuanqian. *Derivation and linear solution of dimensionless IPR curves* [J]. Acta Petrolei Sinica, 2018, 7(2): 63-73.

[15] Zhao Lin, Feng Lianyong, Lu Xiagan, Tong Xiaoguang. *Comparison of two methods for solving the generalized Weng model* [J]. Xinjiang Petroleum Geology, 2017, 30(5): 658-660.