Prediction of shale oil production based on Prophet algorithm

View the article online for updates and enhancements.

Prediction of shale oil production based on Prophet algorithm

Xiaolong Wan, Yongling Zou*, Juan Wang, Weina Wang

Exploration and Development Research Institute of PetroChina Changqing Oilfield Company, Xi'an, China

* Corresponding author: zyl6_cq@petrochina.com.cn

Abstract. The large-scale volume fracturing development of shale oil horizontal wells, and the production is affected by seasonal cycle and emergencies, resulting in the complexity of production prediction. Aiming at the problem of poor actual effect or heavy workload of shale oil production prediction by classical reservoir engineering, reservoir numerical simulation and other methods, a new shale oil production prediction method based on Prophet algorithm is proposed It is easier to find the inherent law from a small amount of data. In this paper, we use the production data of production wells in Huanjiang A reservoir in 2015 to predict the production, and compare the prediction results with long-term and short-term memory neural network (LSTM) and ARPS production decline model. The results show that the prediction accuracy of Prophet algorithm is higher, and it is more accurate for the production of complex shale oil.

1. Introduction

The shale oil field within the source of 1 billion tons in Qingcheng has been found in Ordos Basin. There is no natural industrial production before the reconstruction of tight oil reservoir. Although the standard of industrial production can be met after fracturing, the stable production capacity is limited. Therefore, the shale oil reservoir reconstruction carried out the practice of horizontal well segmented fracturing and horizontal well sectional volume fracturing, which improved the production degree of recoverable reserves. Then, the fine cutting volume fracturing technology test and application of horizontal shale oil wells were carried out, and the output of single well in horizontal well was significantly improved. With the development of shale oil scale, production prediction is also very important. The production prediction is of great significance to understand the reservoir, stable production of oil field and to prepare scientific and reasonable development adjustment plan. However, due to the large-scale fracturing development of horizontal wells, it also causes the complexity of production prediction.

Many scholars use classical reservoir engineering method [3], neural network prediction model [4-6], exponential smoothing method, autoregressive summation moving average model (ARIMA) [7], support vector machine (SVM) [8], reservoir numerical simulation method, long-term and short-term memory neural network (LSTM) [9] to realize the prediction of oil well production. The classical reservoir engineering method has some limitations, such as ARPS production decline model, which is suitable for the reservoir in decline period and has a small application range. The workload of reservoir numerical simulation is large and the timeliness is low. BP neural network model is widely used in production prediction [10], but it does not consider the influence of historical time. Long term and short-term memory neural network (LSTM) is an improved recurrent neural network [11-12], which considers the change trend and correlation of production dynamic data, but it can only make short-term prediction under the influence of cumulative error. Prophet is a new time series prediction model [13-14]. Prophet

provides intuitive and adjustable parameters, which allows analysts to explore different models flexibly in an interactive way, and can simulate multiple seasonal periodic data at the same time [15]. Unlike LSTM and other models, prophet needs a section of historical data close to the point to be measured to predict the point to be measured. By curve fitting, prophet can get the expression of correlation function, which can realize the prediction of any point. Based on the historical data of shale oil, this paper forecasts the shale oil production through prophet, and compares the prediction results with machine learning LSTM neural network method and conventional production prediction method

2. Principle of Prophet algorithm

Prophet algorithm is a differentiated prediction model, that is, the model can be decomposed into trend term, seasonal period term, holiday effect and other components.

$$P(t) = g(t) + s(t) + h(t) + \varepsilon_t \cdot \tag{1}$$

Trend term: there are two functions in prophet algorithm, one is based on logistic regression function, the other is based on piecewise linear function.

Periodic term: time series may contain seasonal trends of various types of periods. Prophet uses Fourier series to simulate the periodicity of time series.

Holiday items: holidays or some major events will have a great impact on time series, and these time points often do not have periodicity. Prophet simulates the impact of each holiday at different time points as an independent model.

3. Data processing and parameter setting

3.1 Data processing

Prophet has its fixed input format, with a total of two columns, DS for time and y for the value of time series [16]. In order to improve the prediction accuracy of the model and eliminate the influence of data dimension, the data should be normalized before simulation.

$$X_i = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

There will be some abnormal data in the original data, and the abnormal points of the actual data should be deleted before the prediction. The advantage of prophet is that even if there is no partial data, the curve fitting will not be affected. This time, we use the following formula to delete the abnormal points.

$$\left| X - X_{avg} \right| > 3 * sd \tag{3}$$

3.2 Parameter setting

The prediction results of the model using default parameters by prophet are not necessarily ideal. Prophet can use manual setting of parameters and automatic optimization by computer to improve the accuracy of model prediction.

Table 1 Parameter setting table

	Parameters	Describe
Trend parameters	growth	"Linear" or "logistic" is used to specify the trend of a linear or logical curve
	changepoints	Include the date list of potential mutation points (if not specified, it will be automatically identified by default)
	n_changepoints	If no mutation point is specified, the number of mutation points to be automatically identified is required. The default is 25
	changepoint_range	The data range of the mutation point is the first 80% of the data set by default
	changepoint_prior_scale	Set the flexibility of automatic mutation point selection, default 0.05
Periodic parameters	yearly_seasonality	The period is the seasonality of the year
	weekly_seasonality	The period is the seasonality of the weak
	daily_seasonality	The period is the seasonality of the day
	seasonality_mode	'additive', additive model and 'multiplicative' multiplicative model

	seasonality_prior_scale	Changing the intensity of the seasonal model, the default value is 10. The larger the value is, the more obvious the periodic rule is
Holiday parameters	holidays	Custom holiday dates, such as spring inspection in March, peak production in June, July and August, etc
	holiday_prior_scale	Change the intensity of the holiday model, the default value is 10, the larger the value is, the more obvious the holiday effect is

1) Select the trend model (Growth). The piecewise linear function and logistic regression function can be selected. When using the logistic regression function, the Capacity value should be set, i.e. Capacity.

2) Change points: by default, prophet will automatically match and learn the turning points so that the prediction curve can predict the turning points properly. If prophet over fits some turning points, it can specify the change points of time series by manual setting to better control the learning and prediction process.

3) periodicity and holiday: the default is the periodicity with year, week and day. You can also set other periods according to the characteristics of the data. The holiday model can specify the corresponding holidays according to the actual mine practice, such as the spring circuit inspection in March and April every year, the peak production period from June to September, etc.

4) Smooth parameter: according to the fitting condition of the model, the changepoint can be adjusted correspondingly_ prior_ Scale is used to control the flexibility of the trend, seasonality_ prior_ Scale is used to control the flexibility of season items, holidays prior scale is used to control the flexibility of holidays.

At the same time, we can use the advantages of computer to optimize the parameters automatically. In the simulation process, the historical data set is divided into training set and test set, and the computer automatically finds the optimal solution to train the model. Then, by continuously adjusting the parameters, and comparing the prediction results with the test set, we can calculate the error, and finally solve the parameter with the minimum error, This parameter is used to predict the future data. The flow chart of production forecast of Prophet algorithm is shown in Figure 1.

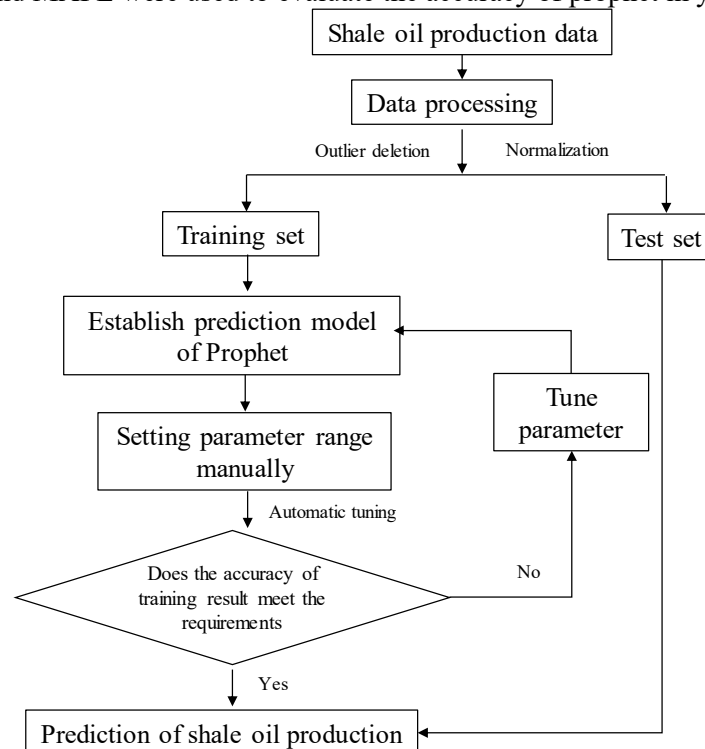Finally, RMSE and MAPE were used to evaluate the accuracy of prophet in yield prediction.

Fig.1 Production forecast process of Prophet

4. Application Examples

Taking Huanjiang a reservoir as an example, the a reservoir mainly develops the turbidite channel microfacies of half deep lake deep lake, and the development layer system is 7. The permeability of the matrix is 0.13md, the porosity is 9.9%, the oil-gas ratio is 95-122m3/t, the average horizontal section length is 1200m, and the well spacing is 300-400M, which is mainly developed by quasi natural energy. This time, the production prediction of 2015 production wells is selected.

4.1 Data processing and parameter setting

This paper uses fbprophet as the learning platform, uses numpy, pandas, pystan and other third-party libraries, and uses python3.5 to write the prediction program.

The data is processed by formula 2 and formula 3. The production data are divided into training set and test set, from July 2015 to December 2019 as training set, and from January 2020 to August 2020 as test set. Combined with the default parameters of prophet, manually set the data range of different parameters (Table 2), where the mutation point n_ Changepoints (25), mutation point range_ Range (80%) selects the default value, and the other parameters are optimized automatically by using the computer according to the set range.

The final parameter is linear; Changepoint_prior_scale = 3; Yearly_seasonality = 5; Seasonality_prior_scale = 10, holidays_prior_scale = 0.5.

Table 2 Different parameter combinations

The parameter name	The parameter value
growth	linear，logistic
changepoint_prior_scale	0.05,0.1,0.3,1,3,10,15,20
seasonality_mode	additive，multiplicative
seasonality_prior_scale	0.01,0.03, 0.1, 0.3, 1, 3, 10,30
holidays	Custom date
holiday_prior_scale	0.05, 0.1, 0.5, 1, 5, 10,15,20

4.2 Production Forecast

The data from the Wells that started production in 2015 were brought into the trained model for production prediction. In order to verify the prediction effect of PROPHET algorithm, the same data is substituted into LSTM algorithm and ARPS production decline model to get corresponding prediction results respectively, and the prediction results of these three algorithms are compared and analyzed and evaluated.

Prophet prediction results were compared with those predicted by LSTM neural network and ARPS production decline model (Fig. 2). Prophet prediction results were significantly better than those predicted by the other two prediction methods.

Table 3 Accuracy comparison of threeprediction methods

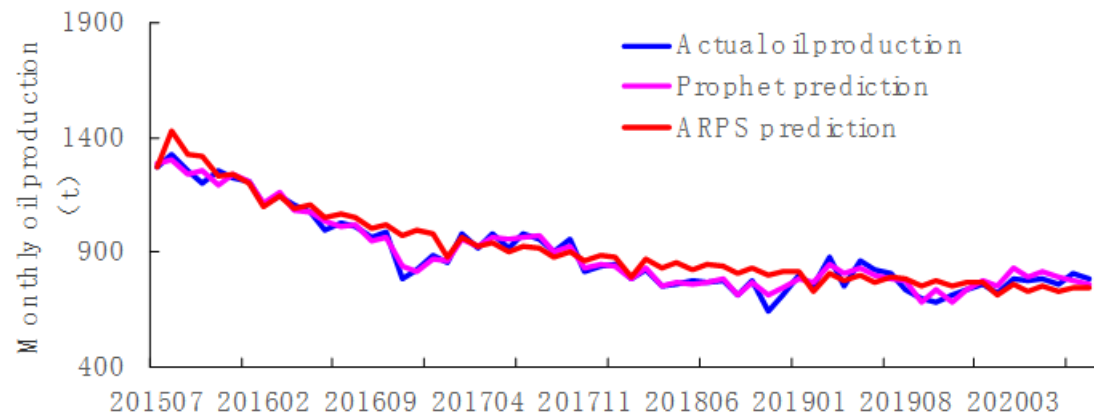| Prediction method	RMSE	MAPE
Prophet predicted	27.00	0.03
LSTM prediction	75.20	0.11
ARPS decline model	58.86	0.05

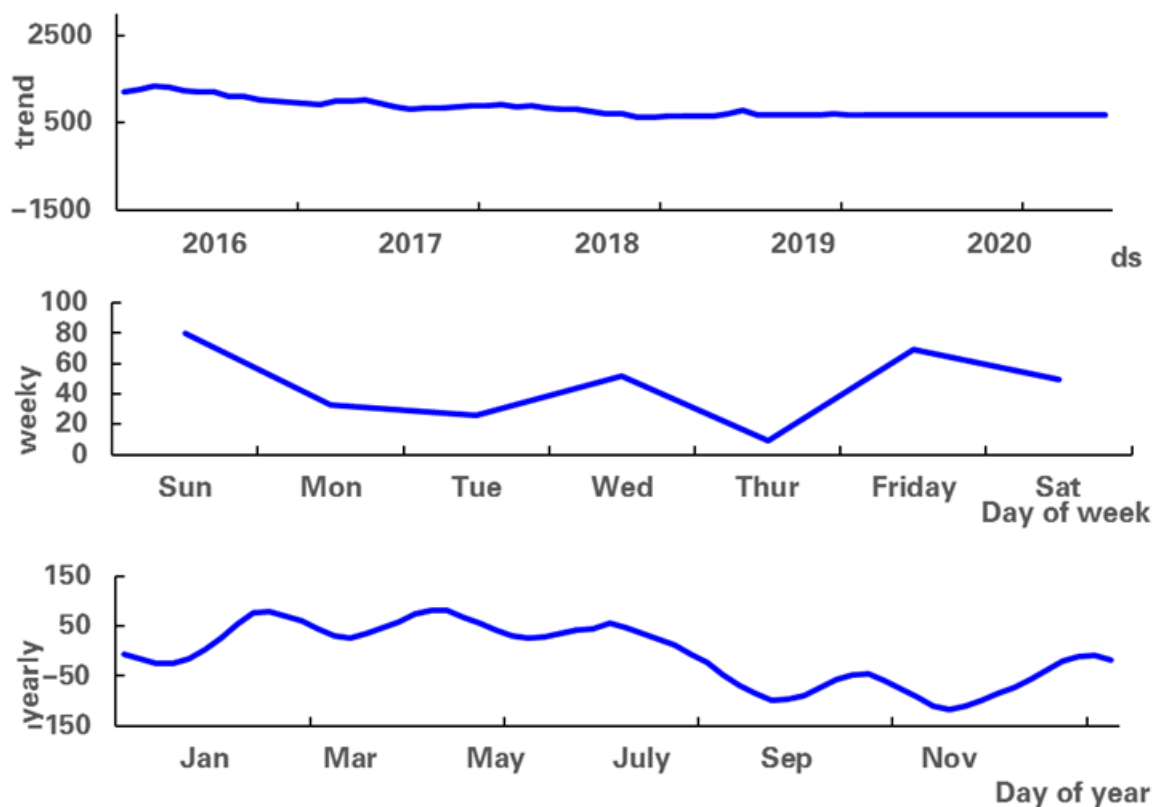Fig. 2 Comparison of shale oil production prediction results

Fig. 3. Composition analysis

As can be seen from the trend item in Figure 3, the yield change trend is relatively obvious, and there is no obvious periodicity. From July 2016 to September 2017, there was an obvious stable yield trend, and after 2018, there was a downward trend. As can be seen from the annual output change chart, there was an obvious peak of output from March to mid-July, which was also consistent with the field cognition. As can be seen from the graph of yield change on a weekly basis, Thursday yield was lower.

The model was used to predict the production of two ultra-low permeability Class III horizontal Wells. The predicted results are in good agreement with the actual situation, which verifies the versatility of this method in the production prediction of fractured horizontal Wells.

5. Conclusion and understanding

Prophet is based on statistics, has complete mathematical theory support, considers holidays, time trend and cycle and other factors, it provides intuitive and adjustable parameters, simple calculation, fast speed, it is easier to find the inherent law of data from a small amount of data, and it can be used for long-term prediction.

Compared with LSTM machine learning method and conventional production prediction method, prophet can realize the medium and long-term prediction of shale oil production. Prophet algorithm enriches the production prediction and provides a certain reference for shale oil production allocation.

Symbol note:

$P(t)$——Predictive value, $g(t)$——Trend term, representing the big trend of data, $s(t)$——Periodic term, representing the periodic change of data, $h(t)$——Holiday item, which represents the impact of special events on data, ε_t——Error term; X_i——Normalized data; X——Original data; X_{min}——Minimum value of original data; X_{max}——Maximum value of original data; X_{avg}——Mean value of original data; sd——standard deviation

References

[1]   FU Jinhua, LI Shixiang, NIU Xiaobing, et al. Geological characteristics and exploration of shale oil in Chang 7 Member of Triassic Yanchang Formation, Ordos Basin, NW China[J]. Petroleum Exploration and Development, 2020, 47(5): 870-883.
[2]   HU Suyun, ZHU Rukai, WU Songtao, et al. Profitable exploration and development of continental tight oil in China[J]. Petroleum Exploration and Development, 2018, 45(4): 737-748.
[3]   LIU Xiaohua, ZOU Chunmei, JIANG Yandong, et al. Basic principle and application of modern production decline analysis[J]. Natural Gas Industry, 2010, 30(5): 50-54.
[4]   LI Chunsheng, TAN Minxi, ZHANG Kejia. Oil well production prediction based on improved BP neural network[J]. Science and Technology and Engineering, 2011, 11(31): 7766-7769.
[5]   TIAN Yapeng, JU Binshan. Shale gas production decline prediction model based on improved BP neural network based on genetic algorithm[J]. Chinese Science and Technology Paper, 2016, 11(15):1710-1715.
[6]   FAN Ling, ZHAO Mengmeng, YIN Chuan, et al. Dynamic analysis method of oil field production based on BP neural network[J]. Fault Block Oil and Gas Field, 2013, 20(2): 204-206.
[7]   YANG Yang, CHENG Yuefei, QIAO Ying, et al. A Study on oil Well Production Prediction Based on Time Series Dynamic Analysis [J]. Journal of Southwest Petroleum University (Science & Technology Edition), 2020, 42(6): 82-88.
[8]   SEEMANN D, WILLIAMSON M, HASAN S. Improving reservoir management through big data technologies [R].SPE 167482, 2013.
[9]   HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
[10]   ZHOU Cailan, LIU Min. Application Research on Oil Production Forecasting Based on BP Neural Network[J]. JOURNAL OF WUHAN UNIVERSITY OF TECHNOLOGY，2009，31(3) : 125－129.
[11]   GU Jianwei, ZHOU Mei, LI Zhitao, et al. Prediction method of oil well production based on data mining based on long-term and short-term memory network model[J]. Special Oil and Gas Reservoirs, 2019, 26(2): 77-81.
[12]   HOU Chunhua. Prediction method of oil production in new wells based on long and short term memory neural network[J]. Oil and Gas Geology and Recovery, 2019, 26(3): 105-110.
[13]   Sean J. Taylor &Benjamin Letham. Forecasting at Scale [J]. The American Statistician,2018,72(1):37-45.
[14]   Taylor S J, Letham B. Prophet: automatic forecasting procedure. R package version 0.3,2017.

[15] Zhao N Z, Liu Y, Jennifer K Vanos et al. Day-of-week and seasonal patterns of PM2.5 concentration over the United States: Time-series analyses using the Prophet procedure[J]. Atmospheric Enviroment,2018,192:116-127.

[16] ZHAI Dulin, ZHANG Xuemin, XIONG Pan, et al. Detection of Ionospheric TEC anomalies based on Prophet Time-series Forecasting Model[J]. EARTHQUCKE, 2019,39（2）：46-62.