STATISTICAL AND NEURAL NETWORK METHODS FOR LOCALIZING RESIDUAL OIL RESERVES

Rinat Safuanov[1]
Rustam Ziniukov[1]
Timur Poryvaev[1]
Timur Petrov[1]
Farhutdinov Ildar[2]

[1] **Russia**, Kazan Federal University, Institute of Geology and Petroleum Technologies
[2] **Russia,** PJSC TATNEFT

ABSTRACT

The determination of residual oil reserves is a complex task that involves high qualified specialists and resource-intensive calculations. Classical oil reservoir simulation methods cannot provide proper level of reliability due to increasing of complexity caused by oil reservoir depleting. Stagnant oil zones arising from heterogeneity of shear modulus of medium cannot be determinate by oil reservoir simulation due to use of effective values of porosity and absolute and phase permeability. The problem of determination of stagnant zones can be solved by using more flexible methods as a machine learning in general and the neural network in particular. The flexibility of these methods allow to build a model that can consider both physical properties of reservoir and the it operating history. In this paper the model based on Convolutional Neural Network (CNN) was built. The choice of CNN based on the assumption that this type of neural network does analysis in a similar way to a human. CNN is also simplifying the approach to distance depended values because the actual coordinate dependence which may not be the continuous function due to border conditions replaces with zone-based dependence produced by convolution. In this case the predicting value function becomes continuous and satisfies Cybenko theorem. The main idea of this model is transforming neighbor wells data of some point in 2D space into **n**x**n** matrix which is similar to grayscale image. The stack of matrices passes to the CNN as different channels of full-color image to find dependencies between image and an actual value which will be predict after successful learning.

Keywords: oil, heavy oil, machine learning, neural network, convolutional neural network

INTRODUCTION

The considered reservoir refers to the Devonian period located in southern part of the Republic of Tatarstan, Russia, represented mainly by heavy oil which contains large amount of heteroatoms compounds. Reservoir generation source is sapropel type organic compounds including humus and aqua humus inclusions with sea type formation [1]. The high viscosity and density promotes the formation of stagnant zones which cannot be determinate easy way. Classical reservoir simulation based on Darcy and Darcy–Forchheimer equations [2] cannot provide proper level of reliability on this type of

https://doi.org/10.5593/sgem2021/1.1/s06.120

reservoir due to complexities caused by reservoir depletion. The solvation of this problem is using different instrument such a neural network [3].

MAIN IDEA

The main idea of the built model is prediction of some value **P** in some point at 2D space of reservoir by analysis of both physical properties and operation history of neighbor well.

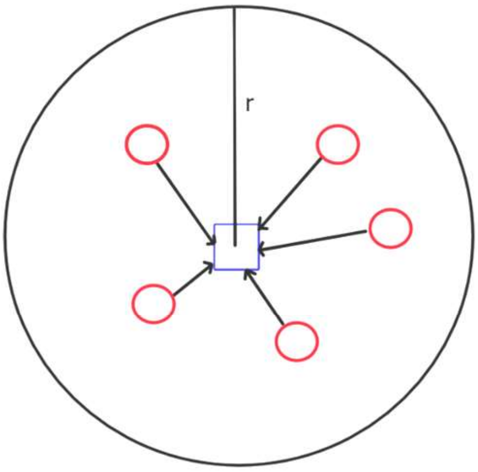

Figure 1. Imaginary point (blue square) in 2D space surrounded by working wells (red circles).

At Fig.1 the imaginary point surrounded by working wells within radius **r** where each well has its own history and physical properties. Due to small value of r (about 500m.) it can be assumed that each well working in the same zone of whole oil reservoir which means that coordinate z is same for all wells. That assumption allows us to make next: if we`ll drill new well in imaginary point location, it work the similar way as its neighbors. So we can predict some value **P** based on neighbors' history and physical properties. The strong definition on the problem is: Predict **P** of the point $\{x_0, y_0\}$ based on neighbor wells data at points $\{ \{x_1, y_1\}, \ldots\}$.

SOLVING THE PROBLEM

Based on assumptions value **P** is a matrix function which transform matrix **m**x**n** where **m** – neighbor wells amount, **n** – amount of passing parameters into single value. According to Cybenko theorem [3], the function **P** should be continuous, but since geometric position of neighbor wells is important, this function may not be continuous due to boundaries conditions caused by heterogeneous nature of oil reservoir. It means that function describing value versus coordinate dependence may not exists at all. This case makes function **P** discontinuous in some areas of oil reservoir. Also **m** is not a constant which means that amount of neighbor wells can vary from 0 (at the reservoir boundaries) to ∞ (some limit amount of wells per km^2). This problem can be solved by using a convolutional neural network.

The zone of radius **r** can be inscribed into square with a side of 2**r**, this square can be divide into small squares within small side **r`**. The given square net can be interpreted as an image where each small square is a pixel, this image can be written as a matrix **K:**

$$\begin{pmatrix} K`_1 & \cdots & K`_n \\ \vdots & \ddots & \vdots \\ K`_{1n} & \cdots & K`_{nn} \end{pmatrix}$$

Each element **K`** represent small square of original square with side a 2r. Value of element **K`** describes by some parameters of well which got into corresponding small square of origin square.

This matrix **K** can be interpret as a grayscale image which may be passed to convolutional neural network [4]. Convolutional neural network consists of two types of layer:

- Convolutional layer – performs convolution of given image with kernel`s matrix which will be determinate in learning process
- Feed-forward layer – summarizes convolution result, performs analysis of summarized values, makes prediction.

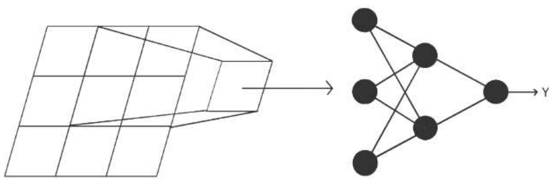

Figure 2. Graphical representation of convolution neural network

By using this type of neural network we solve two problems: the amount of neighbor wells is not a problem anymore and value-coordinate dependence replace with value-zone dependence where zones determinates by size of convolutional kernel size. Value-zone dependence allows different zones to separate dependence from each other. As a pixel is scalar value we need to create and fill the same matrix for each values we want to pass to CNN. Each matrix is the separate channel of one complex image as like Red Green Blue channels of RGB image.

The most common value in all set of **K** matrices will be zero due to absence of wells in the most part of **K`** corresponding squares. This causes uncertainty because wells may be situated in any **K`** corresponding squares and case in which picture of wells location is unique in relation of learning dataset may occurs. In such case the prediction of CNN will be undefined. Since wells amount appropriated for learning is limited we should fill zero values **K`** with some actual data related values. We considered two possible ways:

- Matrix **K** divides into regions that closest to wells. These regions fill by normal or exponential distribution (depends on data nature) with mean as region-owner`s value and standard deviation as standard deviation of all wells value in matrix.
- Matrix **K** fills by linearly extrapolated field based on matrix values.

In first case had observed highly inflated predictions value due to distribution because it may distribute anomaly high value with small chance which increases in large size of dataset. Also the reproducibility of predictions is low because of random nature of distributed values.

The second case is more preferred on considered oil reservoir because of sea type formation. Also it provides high reproducibility due to determine behavior of operation.

Recommended but not necessary operation is smoothing given matrices by Gaussian smoothing, this decreases sharpness of prediction and improve learn rate because of uniformity of input data [5].

UNDERSTANDING OF THE PREDICTION

Based on method described above we can interpret way in which neural network makes prediction: Convolutional neural network watches simultaneously on set of given images,

https://doi.org/10.5593/sgem2021/1.1/s06.120

transforms these into own internal representations by kernels and makes decision based on its own way of “thinking”. It is similar how the real geologist makes his decision – they analyzes the stack of reservoir properties and makes decision based on both data and they experience. The experience part can be integrating by using LSTM neural network, but it goes beyond of this paper.

THE PREDICTION

The **P** not defined till now may be defined as any parameter which could describes residual oil reserves, but the requirement is the correlation between **P** and set of data of **K** matrices. In our work we assume that the oil production rate may be the best thing to describe residual oil reserves because it can be obtained by simple dividing well`s produced oil by its worked time without any complex manipulation, its correlate both with physical properties of reservoir and well`s parameters such as water production, water injection and etc.

In our work we used that set of input parameters of neighbor wells:

- Oil production rate per year– describes how neighbor wells produce oil
- Water production rate per year– describes condition of water encroachment of neighbor wells
- Water injection rate per year – describes both neighbor wells operation mode and amount of water injecting in reservoir
- Well`s pressure – describes pressure at a reservoir depth of neighbor wells
- RQI, Initial oil saturation, Saturated height – describes physical properties of reservoir of neighbor wells
- Cumulative oil, water production, water injection – describes how neighbor wells work in all lifetime
- Length of operating phases – describes how long neighbor wells worked in producing mode, in injecting mode, idle and etc.
- Hydraulic fracturing on commissioning – describes is hydraulic fracturing was carried out upon commissioning because it boosts oil production
- Amount of hydraulic fracturing – describes amount of hydraulic fracturing of neighbor wells
- Oil production per period – describes condition of reservoir for some period
- Influence coefficient of injection wells – describes force of influence of injection wells in neighborhood

RESULTS

The prediction performs on two sites:

The first site consists of 536 wells, 213 were used as learning dataset.

The second site consist of 91 wells, 25 were used as learning dataset.

Convolutional neural network consists of three layers: Input convolutional layer with 34 channels, hidden feed-forward layer within 1300 neurons, output feed-forward layer within 1 neuron.

Training performs with AdamW algorithm [6]. Cost function was MSE.

Training on site 1 took about 5 hours. When cost value decreased to 0.1 training stopped. The model verified by compare real versus predict value for the period 2019-2020:

As can be seen the best match of results observes in range from 1 up to 50 tons per day.

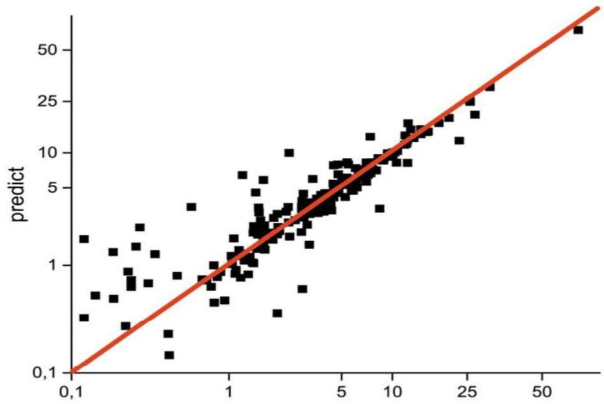

Figure 3 Real Oil production rate vs Predict (ton per day) graph on log coordinates at site 1. Red line shows ideal coincidence between values

The results below 1 ton per day lies chaotically. This can indicate that some important dependences were missed. The Pearson correlation coefficient is 0.84 which says of high reliability of given model on current site.

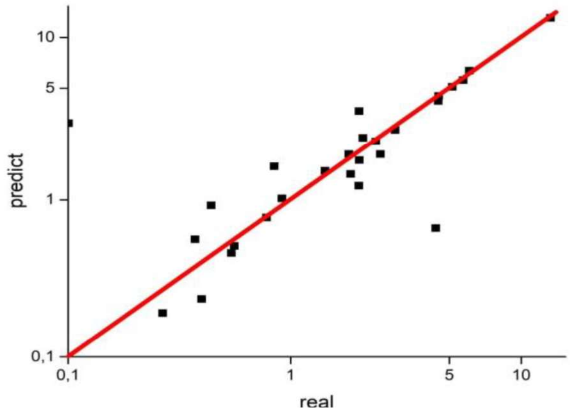

Figure 4. Real Oil production rate vs Predict (ton per day) graph on log coordinates at site 2. Red line shows ideal coincidence between values

Training on site 2 took about 1 hour. When cost value decreased to 0.01 training stopped. Verifying of the model perform same as describes above.

As can be seen, the same problem with range under 1 ton per day does not reproduced. The Pearson correlation coefficient is 0.94 which says of very high reliability of given model on site 2.

After verification prediction maps for each site was built.

Figure 5. One of blocks of prediction map for site 1. Blue crosses mean injection wells, any color dots mean production wells, red and blacks crosses mean well was disabled. Brown color shows value of prediction - the darker the higher

Figure 6. One of blocks of prediction map for site 1. Blue crosses mean injection wells, any color dots mean production wells, red and blacks crosses mean well was disabled. Brown color shows value of prediction - the darker the higher

Fig. 5 shows one block of prediction map for site 1. As can be seen prediction shows that zones with high potential oil production rate located outside of injections wells. Middle of block filled by disabled injection well shows no potential values unlike top right zone which has no wells. It can be interpreting that neural network creates its own analog of "displacement" by injection water. All potential zones about shifted away from central injection well.

On fig. 6 can be seen the same situation as on fig 5. The disabled injection well at bottom left "displaced" prediction values toward to top right where working injection well also "displaces" prediction values in the opposite direction. All prediction value concentrate around disabled production well on top right.

Based on this examples can be said that neural network "understand" the concept of injection water displacement and can handle it in proper way.

The block on fig. 7 shows contradictory situation. The zone around injection well at top side shows usual behavior of "displacing" prediction value as well as two injection wells at bottom left side, but the zone around well under the top injection well shows strange picture of high predicted value line formed. This zone has no reason to show this type of prediction and probably causes by glitch while data generation step.

Fig. 8 shows one block on prediction map from site 2. As can be seen the bottom side of the blocks filled by two disabled wells fill with lighter color than the top left side as well as the top right side unlike the empty top left side. It can be interpreting that neural network "understand" the concept of oil extraction.

Figure 7. One of blocks of prediction map for site 2. Blue crosses mean injection wells, any color dots mean production wells, red and blacks crosses mean well was disabled. Brown color shows value of prediction - the darker the higher

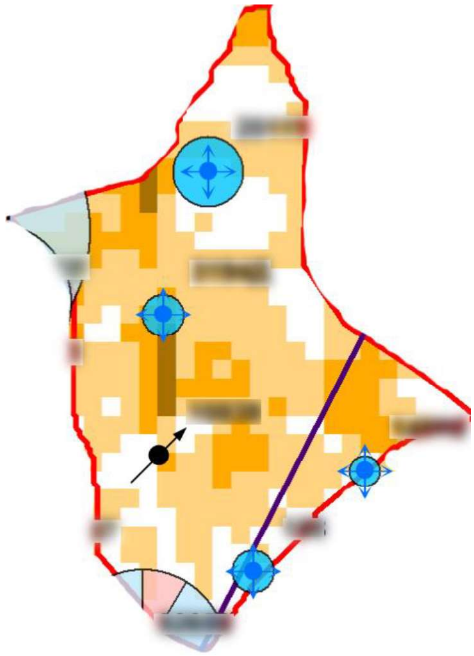

Figure 8. One of blocks of prediction map for site 2. Blue crosses mean injection wells, any color dots mean production wells, red and blacks crosses mean well was disabled. Brown color shows value of prediction - the darker the higher

Based on this example can be said that neural network “understand” not only the “displacing” concept but also the concept of oil extraction. Sometimes neural network can predict in contradictory way which unfortunately can be found only by manual check only.

CONCLUSION

Model was built in this paper base on convolutional neural network shows high reliable result of prediction oil production rate. As shows above transforming neighbour wells data at some point into set of **n**x**n** matrices leads to proper way of determination behavior of some point without knowing any physical properties of it. Described neural network can handle influence of both injection water in reservoir and extraction oil from it. As shows at fig. 5,6,7,8 neural network can “understand” both the concept of “displacing” and the concept of “extraction”. The model passes verifying with high level of confidence. As shows at fig. 3 and 4 the neural network can handle both large and small sites with Pearson correlation coefficient of real vs predict dependence of oil production rate as 0.89 at large and 0.94 at small sites. In paper we assume that oil production rate strongly correlates with residual oil reserves because oil production rate is the most “nature” parameter due to it simple calculation without any errors introduced by it as well as the oil production rate has straight correlation with other production rate parameters such a water production rate, injection rate and etc.

https://doi.org/10.5593/sgem2021/1.1/s06.120

ACKNOWLEDGEMENTS

The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation in accordance with Agreement No. 075-11-2019-032 of 26.11.2019 within the framework of the project " Creation of a high-tech software and hardware complex based on neural network algorithms to improve the efficiency of the development of large hydrocarbon deposits at a late stage"

REFERENCES

[1] Gutorov A. Yu., On the study of the behavior of the variance of the parameters of heterogeneity of productive formations in the development of oil fields in Tatarstan, Oilfield business, Russia, 2015, pp 35-42.

[2] Chavent G., Jérôme J., Mathematical models and finite elements for reservoir simulation: single phase, multiphase and multicomponent flows through porous media, Elsevier, United Kingdom, 1986, pp 10-56.

[3] Cybenko G., Approximation by superpositions of a sigmoidal function, Mathematics of control, signals and systems, USA, 1989, vol. 2, pp 303-316.

[4] O'Shea K., Nash R., An introduction to convolutional neural networks, arXiv preprint arXiv:1511.08458, United Kingdom, 2015, pp 1-10.

[5] Shapiro L. G., Stockman G. C., Computer Vision, USA, 2000, pp 166-176.

[6] Ginsburg, B., Castonguay, P., Hrinchuk, O., Kuchaiev, O., Lavrukhin, V., Leary, R., ... & Cohen, J. M., Stochastic gradient methods with layer-wise adaptive moments for training of deep networks, arXiv preprint arXiv:1905.11286, USA, 2019, pp 1-6.