energies

MDPI

Article

Prediction of Oil Saturation during Water and Gas Injection Using Controllable Convolutional Long Short-Term Memory

Yukun Dong [1,*], Fubin Liu [1], Yu Zhang [1] and Qiong Wu [2]

[1] College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China; z20070088@s.upc.edu.cn (F.L.); z20070066@s.upc.edu.cn (Y.Z.)
[2] Research Institute of Exploration and Development, Tarim Oilfield Company, PetroChina, Korla 841000, China; shhang-tlm@petrochina.com.cn
* Correspondence: dongyk@upc.edu.cn

Abstract: Oil saturation is a kind of spatiotemporal sequence that changes dynamically with time, and it is affected not only by the reservoir properties, but also by the injection–production parameters. When predicting oil saturation during water and gas injection, the influence of time, space and injection–production parameters should be considered. Aiming at this issue, a prediction method based on a controllable convolutional long short-term memory network (Ctrl-CLSTM) is proposed in this paper. The Ctrl-CLSTM is an unsupervised learning model whose input is the previous spatiotemporal sequence together with the controllable factors of corresponding moments, and the output is the sequence to be predicted. In this way, future oil saturation can be generated from the historical context. Concretely, the convolution operation is embedded into each unit to describe the interaction between temporal features and spatial structures of oil saturation, thus the Ctrl-CLSTM realizes the unified modeling of the spatiotemporal features of oil saturation. In addition, a novel control gate structure is introduced in each Ctrl-CLSTM unit to take the injection–production parameters as controllable influencing factors and establish the nonlinear relationship between oil saturation and injection–production parameters according to the coordinates of each well location. Therefore, different oil saturation prediction results can be obtained by changing the injection–production parameters. Finally, experiments on real oilfields show that the Ctrl-CLSTM comprehensively considers the influence of artificial controllable factors such as injection–production parameters, accomplishes accurate prediction of oil saturation with a structure similarity of more than 98% and is more time efficient than reservoir numerical simulation.

Keywords: oil saturation; spatiotemporal sequence; oil recovery; controllable convolutional LSTM

check for updates

Citation: Dong, Y.; Liu, F.; Zhang, Y.; Wu, Q. Prediction of Oil Saturation during Water and Gas Injection Using Controllable Convolutional Long Short-Term Memory. *Energies* **2022**, *15*, 5063. https://doi.org/10.3390/en15145063

Academic Editor: Reza Rezaee

Received: 16 June 2022
Accepted: 9 July 2022
Published: 11 July 2022

1. Introduction

The oil recovery efficiency of an oilfield depends on the geological environment of reservoir and the current level of oil extraction technology. Typically, the oil recovery efficiency is only 30% to 50% in most areas of China, and the remaining 50% to 70% of the oil remains in reservoirs [1]. Therefore, appropriate measures must be implemented to enhance the oil recovery efficiency, including the water and gas injection, which are widely implemented as the secondary oil recovery technology. To improve oil recovery, oilfield experts continuously simulate different production processes by changing injection–production parameters such as water or gas injection rate and water or gas injection timing [2]; sometimes, they also change the well pattern structure, such as adding a water injection well or shutting down a gas injection well. Oil saturation provides a degree of assistance to oilfield experts in their decision-making, which directly represents the oil content in different positions of the reservoir. Clarifying the oil saturation means finding the direction of oilfield development strategy adjustment [3], which has important guiding significance for timely adjustment of the injection–production parameters in the oilfield production site.

Since the mid-20th century, researchers from all over the world have paid extensive attention to the issues related to oil saturation prediction [4]. After continuous research studies, various methods for predicting oil saturation have been proposed, such as seismic analysis, reservoir engineering, well testing and reservoir numerical simulation [5–8]; these traditional methods have made a significant contribution to enhanced reservoir recovery. However, in order to predict oil saturation more quickly and accurately, up-to-date methods should be introduced in the oilfield.

In recent years, with the rapid development of artificial intelligence and big data, the oil and gas resources informatization development has been continuously carried out in the field of petroleum engineering [9]. Those previous conventional studies and the large quantity of accumulated data provide strong supports for deep learning methods to predict oil saturation. Gu et al. used the fully connected long short-term memory [8] (FC-LSTM) network to predict oil saturation, and the influence of different injection–production working systems on oil saturation was considered in the training stage. FC-LSTM handles the temporality of the data naturally, but cannot model the spatial structure of the reservoir sufficiently. However, both temporality and spatiality should be taken into account when building neural networks [10]. Zhang et al. used a convolutional neural-network-based model to predict oil saturation maps [11], which considered various reservoir properties including permeability and active grid. They transformed each of these reservoir properties into a two-dimensional array and made these arrays into a stacked multilayer array, then they used this stacked multilayer array to generate oil saturation maps. Nevertheless, convolutional neural networks could not learn the temporal features effectively, so the whole training process did not model the temporal features of oilfield production processes.

In fact, the treatment of spatiotemporal features is a general problem, which has already been researched by existing studies. ConvLSTM [12], a convolution long short-term memory network, embeds a convolution operation into the LSTM network and replaces the full connection layer with convolution; it obtained good results in many kinds of spatiotemporal sequence prediction tasks [13]. In ConvLSTM, the memory state is only transmitted in the temporal direction, so the spatial modeling capability can still be further improved [14,15]. Based on the prediction method of the spatiotemporal long short-term memory network ST-LSTM [16], a longitudinal spatial memory state transition path was introduced into the network, which effectively reduced the gradient disappearance between neurons in longitudinal layers, so that the final prediction results could widely consider the hierarchical spatiotemporal features.

Oil saturation data consist of special spatiotemporal sequence data, which need to consider simultaneously the effect of geological properties and injection–production parameters [17]. On this basis, this paper proposes a prediction method based on a Ctrl-CLSTM, which adopts a dual-memory state transition mechanism to achieve the unified modeling of the temporal and spatial characteristics of oil saturation. Moreover, we primarily consider the effect of artificial and operable conditions on the oil saturation. By adding a new control gate structure to the network, the model learns more about the influence of the injection–production parameters on oil saturation, which improves the prediction accuracy and makes the prediction process more in line with physical laws. Compared with the traditional numerical simulation, the Ctrl-CLSTM adjusts the injection–production parameters more quickly and maintains an equivalent prediction accuracy. Analyzing the oil saturation generated from different combinations of injection–production parameters, a theoretical reference can be provided to oilfield experts when designing injection–production working systems.

2. Spatiotemporal Sequence Prediction

A spatiotemporal sequence describes the multidimensional grid information, which is applied to various spatiotemporal sequence prediction tasks, such as short-term precipitation prediction, traffic flow prediction, marine pollution trajectory prediction, etc. In addition to changing with time, a spatiotemporal sequence is also interdependent between adjacent spatial grids [18]. The oil saturation is also a multidimensional grid sequence that changes dynamically with time, and the distribution states of adjacent spatial locations affect each other, so it can be defined as a spatiotemporal sequence, and its temporal and spatial characteristics must be learned simultaneously.

Our purpose is to capture the complex dynamic physical laws from the spatiotemporal system of oil saturation continuously observed, and approximately estimate the future trend of oil saturation. A generalized approach for predicting oil saturation using deep learning can be summarized below: the first is to create a training sample set through the reservoir numerical simulation software tool and obtain a large number of grid data of size $m \times n$ with spatial correlation. For a multilayer reservoir, each training sample contains r oil saturation maps; then, the oil saturation at time t can be recorded as $X_t \in R^{m \times n \times r} (r = 1, 2, 3, \ldots)$. Let $X_{\alpha:\beta}$ be the oil saturation sequence of time period $\alpha \sim \beta$, in Equation (1), the sequence $X_{\alpha:\beta}$ is the input of the neural network, and the sequence $\hat{X}_{\beta:\gamma}$ is the most approximate estimation of the future oil saturation sequence $X_{\beta:\gamma}$.

$$\hat{X}_{\beta:\gamma} = \underset{X_{\beta:\gamma} \in R^{m \times n \times r}}{argmax} \; p(X_{\beta:\gamma} | X_{\alpha:\beta}) \tag{1}$$

Of note, the inherent characteristics of a reservoir, such as permeability distribution, formation thickness, boundary dip angle and other factors, have a great impact on the distribution of oil saturation. Moreover, in order to improve oil recovery, oilfield experts usually design a variety of well locations, as well as injection–production working systems, including different parameters such as displacement agents, injection amounts and injection timing. Changing the value of the injection–production parameters will also affect the state of the crude oil in the reservoir; for example, increasing the bottom pressure will make the crude oil have a stronger flow capacity in the reservoir pores [19]. In summary, oil saturation data are special spatiotemporal sequence data, which are affected by both reservoir environment and artificial controllable injection–production working systems in the development process. It is generally believed that the physical property parameters of a reservoir cannot be controlled by humans; in view of this, the change of oil saturation caused by injection–production working systems was considered mainly. The composition of the injection–production working system P is shown in Equation (2). Although the effect of reservoir parameters was not considered much, the Ctrl-CLSTM is an unsupervised learning model and its input contains historical oil saturation, from which reliable potential reservoir properties information can be learned.

$$P = [T_1, \ldots, T_n; G_1, \ldots, G_n; W_1, \ldots, W_x; Q_1, \ldots, Q_y] \tag{2}$$

where n, x and y represent the number of gas injection wells, water injection wells and production wells, respectively, T_i and G_i represent the gas injection timing and gas injection speed, respectively, W_i indicates the water injection rate and Q_i indicates the bottom hole pressure of the production wells $(i = 1, 2, 3, \ldots)$.

In order to ensure the value of each injection–production parameter evenly covered the entire value range, the Latin hypercube sampling [20] was used to generate the experimental data. Some examples of injection–production parameters of the Tarim oilfield in China are shown in Table 1, where each row represents an injection–production working system that lasts for ten years, and the results files are generated by numerical simulation software semiannually.

For instance, in the first row, T_1 = 3.5 means that the gas injection well G_1 is opened in the 3.5th year, and its gas injection rate is 93,459 m^3/d, W_1 = 124 indicates that the water injection volume of water injection well W_1 is 124 m^3/d, and Q_1 = 54 indicates that the bottom hole flow pressure of production well Q_1 is set to 54 MPa.

Table 1. Partial injection–production parameters.

T_1	G_1	…	T_4	G_4	W_1	…	W_{12}	Q_1	…	Q_{21}
3.5	93,459	…	9	2491	124	…	120	54	…	37
0	113,668	…	8	3679	143	…	78	101	…	70
2	44,748	…	3	6815	21	…	68	123	…	117
6.5	1045	…	0.5	101,057	27	…	76	85	…	33
8	62,062	…	9.5	76,505	67	…	40	74	…	85
1.5	58,784	…	7.5	29,220	52	…	31	70	…	30
0	25,807	…	7.5	116,092	126	…	122	54	…	28
4.5	33,229	…	5	114,720	132	…	144	78	…	80

3. Controllable Convolutional Long Short-Term Memory Network

The Ctrl-CLSTM constructed in this paper takes the historical spatiotemporal sequence and the corresponding controllable factors as the inputs and predicts the future spatiotemporal sequence. We innovatively introduced controllable factors in the spatiotemporal sequence prediction task to specifically deal with a special sequence such as oil saturation. The core part of the Ctrl-CLSTM is the dual-memory state transmit mechanism, which realizes the unified modeling of the spatiotemporal features. In addition, the control gate structure is embedded in all nodes of the neural network, which enhances the modeling ability of the network for controllable factors such as well location distribution, injection–production parameters, etc., controlling the change of the spatiotemporal sequence effectively.

3.1. Architecture of Ctrl-CLSTM

For spatiotemporal sequence prediction tasks, different types of features are encoded into different networks architectures. LSTM is a variant of the recurrent neural network, which is suitable for sequence data prediction with strong temporal dynamics. Its transmission from input state to the hidden state uses the matrix multiplication of the fully connected layer, and the feature state is transferred in the time dimension. A convolutional neural network performs better in dealing with grid data with spatial correlation [21]. ConvLSTM [12] replaces the full connection layer in LSTM units with a convolution layer, which solves the problem that traditional LSTM cannot describe spatial structure to a certain extent. However, the memory state of ConvLSTM is only transmitted along the temporal direction, and the neurons between the vertical layers are independent of each other. The spatial features of ConvLSTM become abstract with the depth of the convolution operation, and the ability to maintain the detailed features is not enough. On this basis, ST-LSTM [16] adds a spatial memory state transmitted between longitudinal layers, which effectively reduces the gradient disappearance of spatial features and alleviates its abstraction. Finally, ST-LSTM deeply fuses the temporal memory state and spatial memory state to learn hierarchical spatiotemporal features widely. The Ctrl-CLSTM adopts this dual-memory state-transfer mechanism like ST-LSTM, and innovatively embeds the control gate structure in each neural unit, making the controllable factors as an additional input.

The network structure of the Ctrl-CLSTM is shown in Figure 1, where l represents the number of layers, t represents the time, the bottom layer is the input unit and the top layer is the output unit. The horizontal transfer path encodes the temporal state C_t^l, with arrows pointing to the next moment, and the vertical transfer path encodes the spatial state M_t^l, with arrows pointing to a deeper level of feature abstraction. The spatial state of the vertical transfer path is transmitted to the bottom unit of the next moment after reaching the top unit, which effectively reduces the gradient vanishing phenomenon. Apart from that, we can learn the memory sates C_t^l and M_t^l from memory networks, in which they can be seen as additional computational steps [22].

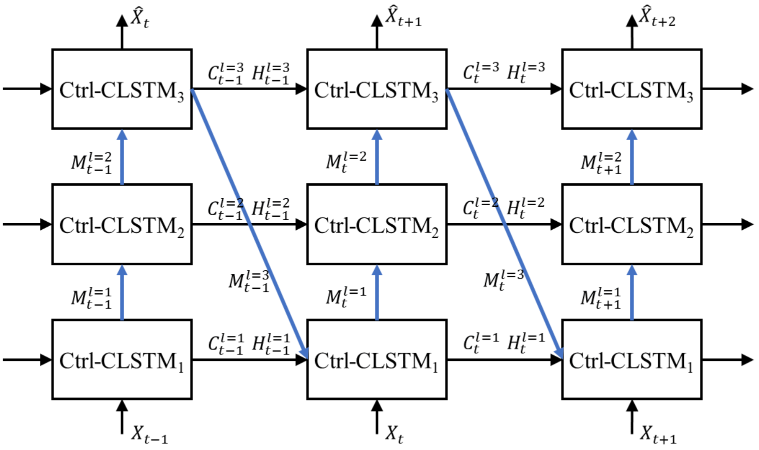


Figure 1. Architecture of Ctrl-CLSTM.

3.2. Units of Ctrl-CLSTM

The role of controllable factors is to intervene in the evolution of a spatiotemporal sequence and affect its development direction; different controllable factors contribute to different degrees of prediction results. As mentioned in Section 1, oil saturation is affected by many factors, especially in reservoirs with strong heterogeneity. Due to uneven distribution of oil layers and large differences in reservoir permeability, the influence of the injection–production parameters is particularly prominent [23,24]. Therefore, we select the injection–production parameters as the primary controllable factor. By introducing the control gate structure, the Ctrl-CLSTM learns the dynamic mechanism of injection–production parameters to a certain extent, integrates injection–production parameters and oil saturation and extracts their characteristics to capture the influence of the injection–production parameters on oil saturation changes.

As shown in Figure 2, the state transition path of a Ctrl-CLSTM unit is divided into two parts, including the temporal memory state C_t^l (Blue) and the spatial memory state M_t^l (Orange). In the stage of data preprocessing, the injection–production parameters should be combined with the well location map, and the geographical spatial position of the working wells is crucial to the production. Therefore, the control gate structure $k_t^{'}$ (yellow) is embedded in the spatial memory state transition path. This control gate structure increases the number of nonlinear neurons involved in the state transition calculation in each unit node, which is equivalent to expanding the memory state transition path in the spatial dimension, and conducive to improving the correlation of memory states at different spatial layers. At the same time, it enhances the modeling ability of the model for injection–production parameters and well location map, so that the network can learn the control effect of the injection–production parameters on oil saturation. To date, there are four gate structures on the M_t^l path, that is, the input gate $i_t^{'}$, which selectively retains the input passed in from the previous neuron, focusing on recording the important information; the renew gate $g_t^{'}$, which determines the content to be written; the forget gate $f_t^{'}$, which selectively forgets the input passed in from the previous neuron, forgetting the unimportant information; and the novel control gate $k_t^{'}$. In this way, it can be distinguished from the control signal of the temporal memory state C_t^l. The specific state transition method of the Ctrl-CLSTM is as follows.

Temporal memory state transition path:

$$\begin{aligned}
g_t &= tanh(W_g * (S_t \oplus P_t) + W_{hg} * H_{t-1}^l) \\
i_t &= \sigma(W_i * (S_t \oplus P_t) + W_{hi} * H_{t-1}^l) \\
f_t &= \sigma(W_f * (S_t \oplus P_t) + W_{hf} * H_{t-1}^l) \\
C_t^l &= f_t \odot C_{t-1}^l + i_t \odot g_t
\end{aligned} \tag{3}$$

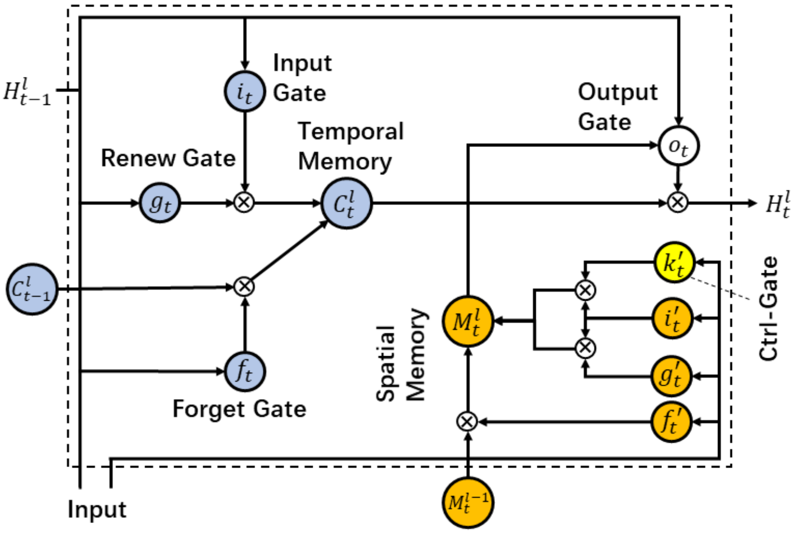


Figure 2. Units of Ctrl-CLSTM.

Spatial memory state transition path:

$$\begin{aligned} g_t' &= tanh(W_g' * (S_t \oplus P_t) + W_{mg} * M_t^{l-1}) \\ i_t' &= \sigma(W_i' * (S_t \oplus P_t) + W_{mi} * M_t^{l-1}) \\ f_t' &= \sigma(W_f' * (S_t \oplus P_t) + W_{mf} * M_t^{l-1}) \\ k_t' &= tanh(W_k' * (S_t \oplus P_t) + W_{mk} * M_t^{l-1}) \\ M_t^l &= f_t' \odot M_t^{l-1} + i_t' \odot g_t' + i_t' \odot k_t' \end{aligned} \tag{4}$$

The output gate o_t further seamlessly fuses the control signals from both directions:

$$o_t = \sigma(W_o * (S_t \oplus P_t) + W_{ho} * H_{t-1}^l + W_{co} * C_t^l + W_{mo} * M_t^l) \tag{5}$$

Equations (3)–(5) are the state transition path of the Ctrl-CLSTM, where S_t is the oil saturation at time t, P_t is the injection–production parameters corresponding to time t and $\oplus$ links S_t and P_t in parallel. They establish a matrix with the same size as the oil saturation, then mark the water injection wells, gas injection wells and production wells on the matrix according to the well position coordinates; afterwards, they record the production parameters of the corresponding wells, and finally get a dual-channel matrix in parallel with the oil saturation. $(S_t \oplus P_t)$ is the input of the Ctrl-CLSTM, including the oil saturation, well location information and injection–production parameters. $*$ means the convolution operation, $\odot$ is the Hadamard product [25], which is the multiplication of the corresponding elements of two matrices of the same order. σ represents a sigmoid activation function, *tanh* is a nonlinear hyperbolic activation function, both of which are applied between neurons in two adjacent layers, making the neural network have nonlinear mapping capability. $W_{(.)}$ is the connection weight corresponding to each Ctrl-CLSTM gate.

$$H_t^l = o_t \odot tanh(W_{(1\times1)} * [C_t^l, M_t^l]) \tag{6}$$

Hidden state H_t^l is the output of a Ctrl-CLSTM unit at time t and depends on the nonlinear combination of C_t^l and M_t^l. Equation (6) reduces the dimension of the hidden state through a 1×1 convolution layer and a *tanh* activation function, resulting in H_t^l having the same dimension as C_t^l and M_t^l. In particular, the bottom unit receives the memory state of the top unit at the previous time: $M_t^{l=0} = M_{t-1}^{l=3}$.

As shown in Figure 1, the temporal memory state C_t^l is transmitted horizontally, the spatial memory state M_t^l is transmitted longitudinally and the whole is transmitted in a zigzag manner. In order to make the temporal memory state and the spatial memory state update their nonlinear functions independently at any time, the two should remain orthogonal during training. Therefore, in Equation (7), a decoupling loss function $L_{decouple}$ is defined through the cosine similarity function, so that the two transmit their own characteristic parameters in their respective latent spaces independently without interfering with each other, where $W_{decouple}$ represents the convolution kernel of 1×1, $\langle , \rangle_c$ is the matrix point multiplication, and $\|.\|_c$ represents the L2 regularization operation after flattening the features, which is calculated according to the channel c.

$$\begin{aligned} \Delta C_t^l &= W_{decouple} * (i_t \odot g_t) \\ \Delta M_t^l &= W_{decouple} * (i_t' \odot g_t' + i_t' \odot k_t') \\ L_{decouple} &= \sum_t \sum_l \sum_c \frac{|\langle \Delta C_t^l, \Delta M_t^l \rangle_c|}{\|\Delta C_t^l\|_c \cdot \|\Delta M_t^l\|_c} \end{aligned} \tag{7}$$

4. Experiments and Results

In this paper, an actual reservoir officially developed in the Tarim oilfield since 1999 was selected as the research object. The data such as water content, gas-oil ratio and static pressure have a good history matching. The reservoir model consists of 471 × 421 × 7 grids, and each grid is 10 m × 10 m × 10 m in size. There are 21 production wells and 16 injection wells in this reservoir, 12 of which are water injection wells and the other 4 are gas injection wells. Figure 3 shows the locations of the wells on the permeability distribution map, where 4 triangles mark the locations of gas injection wells, 12 pentagrams mark the locations of water injection wells, and 21 dots indicate the locations of production wells. We simulated 15 well pattern structures, each with 300 injection–production working systems, for a total of 4500 injection–production working systems. As for the value of injection–production parameters, we used the Latin hypercube sampling method to sample uniformly within a certain range. The bottom hole pressure of a production well was at least 28 MPa, the injection volume of a water injection well was in the range of 0–150 m^3/d and the injection volume of a gas injection well was in the range of 0–120,000 m^3/d. Finally, all injection–production working systems were fitted by the numerical simulation technology as samples. The number of samples was sufficient, which was suitable for neural network training and learning.

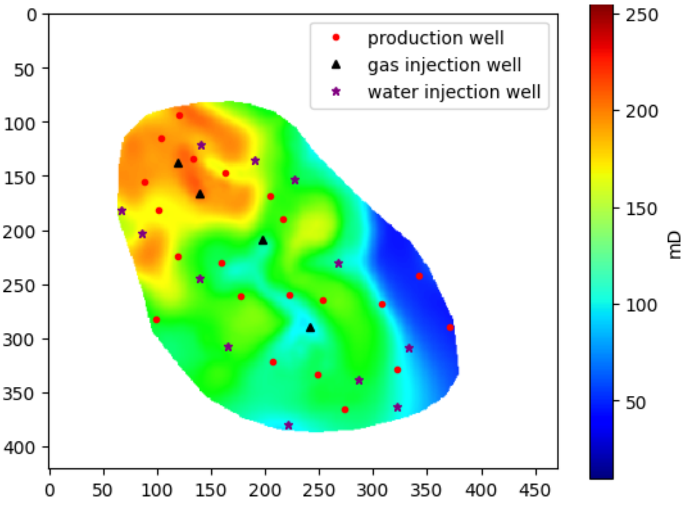


Figure 3. Permeability distribution and well location.

4.1. Constructing the Dataset

Making the data suitable for our model required the following steps:

(1) Latin hypercube sampling (LHS): Sampling is often present in experiments, and a good sampling strategy can make our model's sample point distribution more reasonable. We adopted LHS to generate uniformly distributed injection–production parameters. LHS is a stratified random sampling that allows efficient sampling from the distribution interval of the variables, ensuring a comprehensive coverage of the value range of each injection–production parameter. A LHS X_n in [0, 1) is defined by the following equation:

$$x_{n,j} = \frac{\pi_j(n-1) + U_{n,j}}{N} \tag{8}$$

Among Equation (8), $x_{n,j}$ is the jth component of sample point X_n, π_j is a permutation of $[0, N-1]$, uniformly chosen at random, N is the total number of X_n and $U_{n,j}$ is uniformly distributed in [0, 1).

(2) We adjusted the oil saturation data generated by the numerical simulation to a matrix of size 471 × 421. Then, Equation (9) normalized the injection–production parameters to reduce the dimensional gap between different well data and help to reduce the loss function value, where x is the parameter to be normalized, x_{max} and x_{min} are the maximum and minimum parameters of this working system and $x^{'}$ is the result after normalization.

$$x^{'} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{9}$$

(3) In order to utilize the well location information comprehensively, we marked and distinguish the gas injection wells, water injection wells and production wells on the well location map according to the well location coordinates, corresponding to the positions of the triangles, pentagrams and dots in Figure 3. The injection–production parameters were written at the corresponding coordinates, then we linked well location map and oil saturation in parallel to form a dual-channel matrix. In order to balance the training duration and the prediction accuracy, we discarded the area around the oilfield where the oil content was always zero, and only the effective values in the matrix were retained during the training process. Finally, the size of the dual-channel matrix was adjusted to 256 × 256 × 2, which reduced the occupation rate of the training display RAM and sped up the convergence.

(4) According to the method proposed by Wang [15,16], we regarded the 10 parallel dual-channel matrices as an entire spatiotemporal sequence, including five historical observations (input of the model) and five future prediction results (output of the model), to avoid data interference between different working systems.

(5) Dataset division: we selected 200 groups of injection–production working systems from every well pattern structure as the training set, 50 groups as the validation set, and 50 groups as the test set, which evenly covered all types of well pattern structures.

4.2. Implementation Details

We observed that the number of layers of the Ctrl-CLSTM had a great influence on the final performance. After repeated trials, we finally set three stacked layers in the Ctrl-CLSTM, with 128 channels for every memory state, to maintain a balance between the training efficiency and prediction quality. Unless otherwise specified, we used the ADAM optimizer [26] to train the model, which is one of the most common optimization algorithms and has been validated in a large number of deep neural networks. The initial learning rate was set to 0.001 and the convolution kernel size was set to 5 × 5. Equation (10) is the loss function used, which is the sum of the reconstruction loss function and the decoupling loss function; the deviation of the generated image from the real image is defined by the reconstruction loss function, where $\hat{X}_t$ and X_t are the predicted oil saturation and the actual image, respectively, in each time step t, $||.||_2^2$ is the square operation and λ is a constant.

$$Loss = \sum_{t=\beta}^{t=\gamma} ||\hat{X}_t - X_t||_2^2 + \lambda L_{decouple} \tag{10}$$

We also noted that the selection of batch size had a great impact on the convergence speed of the loss function; therefore, we selected a minibatch of four, that is, each training iteration trained four sequences randomly at once, and the training was stopped after 50,000 training iterations. All intelligent models were set with similar hyperparameters and the experimental code was written in PyTorch, which ran on a GeForce RTX™ 3090 graphics card.

4.3. Evaluation Metrics

We selected the evaluation metrics widely adopted in the tasks of spatiotemporal sequence prediction to evaluate the trained models: the structural similarity index measure (SSIM) [27], evaluating the similarity between the predicted results and the real pictures (the value range is (0, 1) and higher is better).

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{11}$$

Among Equation (11), μ_x and μ_y are the mean intensity of the real picture x and the predicted result y, σ_{xy} is the covariance of x and y, σ_x and σ_y are their standard deviations and C_1 and C_2 are the constant to avoid a zero denominator.

The second metric was the mean squared error (MSE), estimating the absolute pixel level error of two images (lower is better); in Equation (12), m and n are the dimensions of the images and x_{ij} and y_{ij} are the pixel values of the real image and the predicted image, respectively.

$$MSE = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - y_{ij})^2 \tag{12}$$

The third metric was the peak signal-to-noise ratio (PSNR) [28], evaluating the pixel-wise similarity between the predicted image and the real image (higher is better). Generally speaking, a PSNR over 40 dB means that the picture has a low level of distortion.

$$PSNR = 10\, log_{10}\, (255^2 / MSE(x, y)) \tag{13}$$

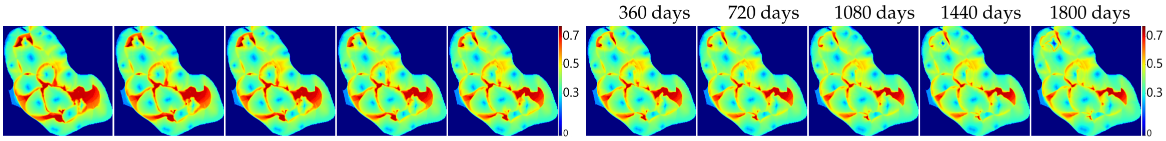
PSNR is one of the most widely used objective evaluation index of images, but it does not consider the visual recognition and perception characteristics of the human eyes, and the evaluation results are often different from the subjective perception of human beings. So we introduced the learned perceptual image patch similarity (LPIPS) [29], which aligns better with human perceptions (lower is better).

$$LPIPS = \sum_l \frac{1}{H_l W_l} \sum_{h,w} ||w_l \odot (x_{hw}^l - y_{hw}^l)||_2^2 \tag{14}$$

In Equation (14), Zhang et al. [29] extracted feature stacks x_{hw}^l and y_{hw}^l from L layers and unit-normalized them in the channel dimension. The vector w_l was used to scale the activations channel-wise and H, W are the size of the feature stack.

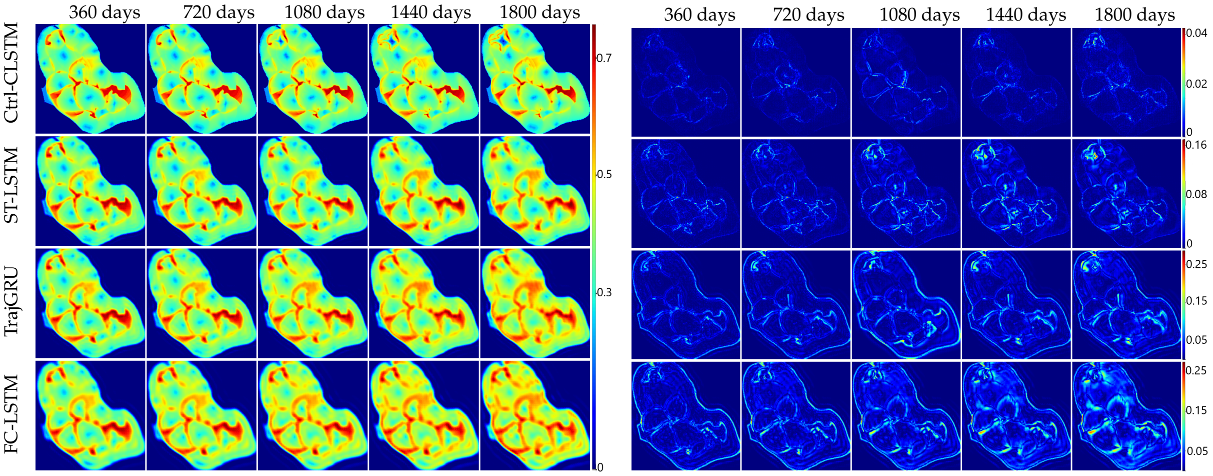
4.4. Comparative Analysis of Oil Saturation Prediction Results of Different Networks

In this section, we compare and analyze the oil saturation predicted by different networks. The input of each model consisted of the previous oil saturation sequence (the sequence in Figure 4a), the injection–production parameters and the well location distribution of a corresponding period. Figure 4b shows the numerical simulation results, which is regarded as the ground truth $\hat{X}_t$ for comparison with other advanced spatiotemporal sequence prediction models, including the traditional long short-term memory network FC-LSTM [8], the TrajGRU [14], which adopts an encoding–predicting network structure able to learn subnetwork parameters and improve the overall network state, and the ST-LSTM [16], which introduces spatial memory flow. Some experimental results of each model on the test set are presented in Figure 5. Each row of Figure 5a shows the predicted oil saturation sequence X_t of the Ctrl-CLSTM, ST-LSTM, TrajGRU and FC-LSTM, for a total of five oil saturation distributions at 360, 720, 1080, 1440 and 1800 days. Figure 5b is also from day 360 to day 1800, and it shows the error ($\hat{X}_t - X_t$) between the numerical simulation result and the prediction of each model.



(**a**) Input sequence (**b**) Numerical simulation results

Figure 4. Input sequence and numerical simulation sequence.



(**a**) Predicted oil saturation at different times (**b**) Error ($\hat{X}_t - X_t$) at different times

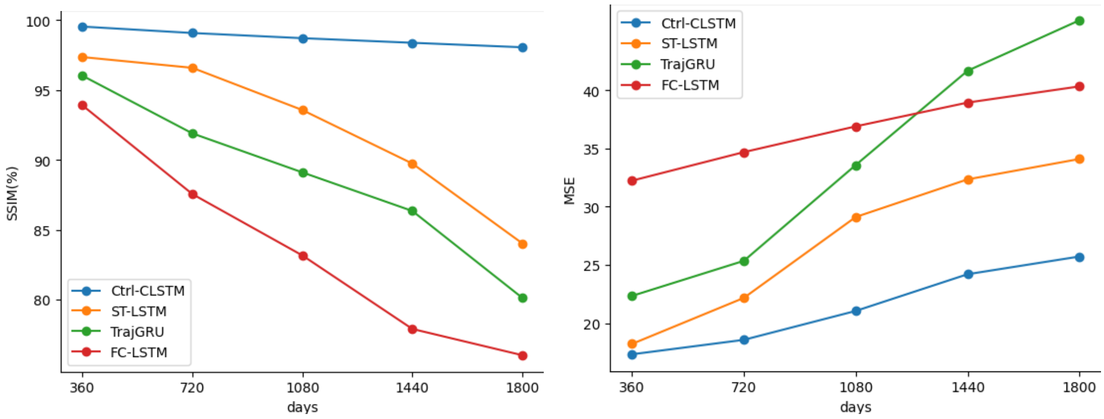
Figure 5. Prediction examples on different models.

It is evident in Figure 5 that the oil saturation predicted by the Ctrl-CLSTM has the highest clarity, and the absolute value of the error is less than 0.04. As the time step of the forecast grows, the performance of other models becomes progressively worse, especially in the boundary where the oil content varies significantly. All of them indicate that, by embedding the control gate structure in each unit, the Ctrl-CLSTM has better learning capability and simulates the changing process of oil saturation realistically.

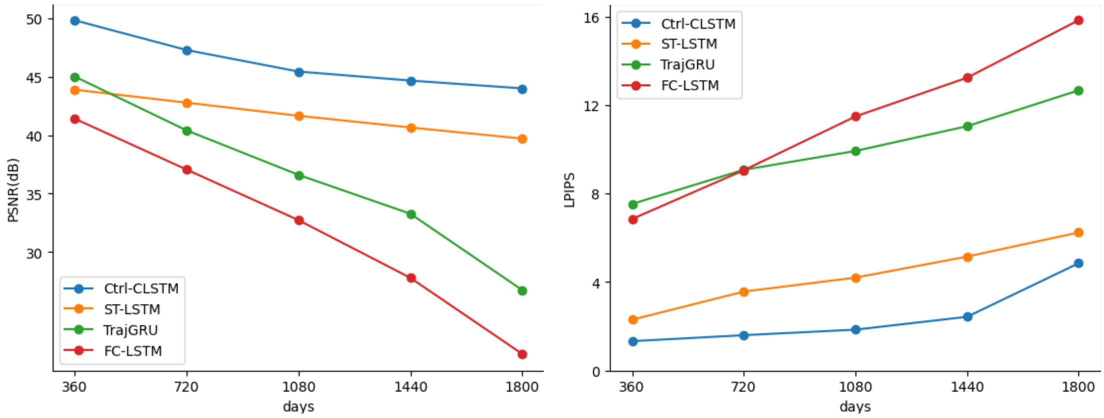
The average quantitative results on each evaluation metric of different models were calculated and are displayed in Table 2, from which we can draw some conclusions: the Ctrl-CLSTM has the best performance in each metric, with the SSIM exceeding 98% and the MSE decreasing from 36.61 to 21.37. LPIPS is closer to human perceptual judgment, and the Ctrl-CLSTM has minimal LPIPS indicating it has excellent power to generate high-fidelity images. With the proposed control gate structure, there is a more accurate grasp of the effect of injection–production parameters on oil saturation. Figure 6 provides the comparison results of corresponding times; as the prediction time step increases, although all models tend to deteriorate, the Ctrl-CLSTM deteriorates more slowly, and the final prediction remains highly accurate and outperforms other models.

Table 2. Quantitative results on each evaluation metric.

Model	SSIM(%)	MSE	LPIPS	PSNR(dB)
FC-LSTM	83.16	36.61	11.29	32.06
TrajGRU	89.10	33.78	10.04	36.41
ST-LSTM	93.56	27.19	4.29	41.73
Ctrl-CLSTM	**98.23**	**21.37**	**2.41**	**46.23**



(**a**) Trend of SSIM over time. (**b**) Trend of MSE over time.



(**c**) Trend of PSNR over time. (**d**) Trend of LPIPS over time.

Figure 6. Prediction effects over time.

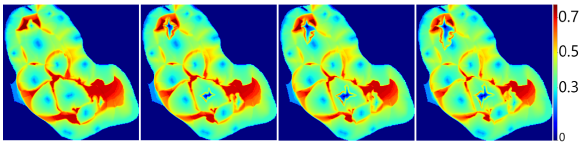
When the Ctrl-CLSTM is applied to predict oil saturation sequence, an appropriate number of model parameters should be chosen. To obtain satisfactory experimental results, the neural network was set to two, three, and four layers separately, and the training results are shown in Table 3. It can be seen that the model performs best when three layers of Ctrl-CLSTM units are designed. When the model uses two stacked layers, the prediction effect is the worst, because only two layers of neurons can not remember the long-term variation of oil saturation. When the model uses a four-layer stack structure, the effect is also reduced, because there are too many neural network parameters at this time, which is prone to overfitting, resulting in a better effect on the training set, but a worse effect on the test set.

Table 3. Quantitative results with different layers.

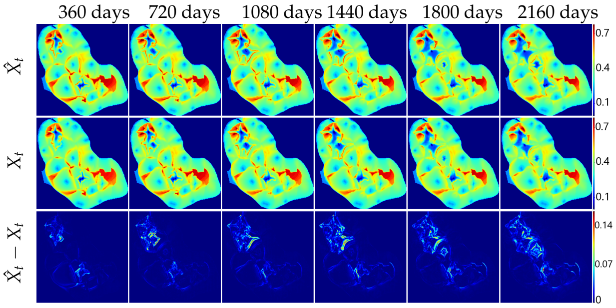
Model Layers	SSIM (%)	MSE	LPIPS	PSNR (dB)
Two layers	90.13	27.82	6.24	42.08
Three layers	**98.23**	**21.37**	**2.41**	**46.23**
Four layers	92.07	25.61	8.93	39.75

4.5. Results on Different Injection–Production Parameters

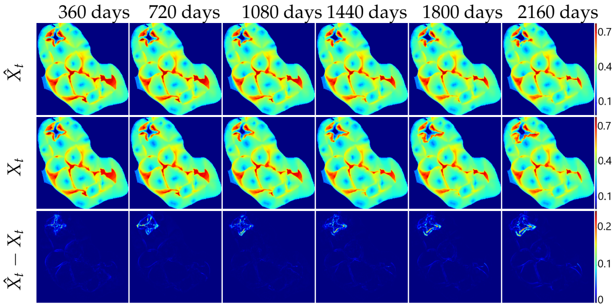
In the case of oil–gas–water three-phase seepage, it is more complicated to formulate a reasonable injection–production working system. Through the control gate, the dynamic physical laws between injection–production parameters and oil saturation can be captured. To verify this point of view, we took the same oil saturation as input sequence, only adjusting the injection–production parameters, then we predicted and observed the variations of oil saturation in the future. The sequence in Figure 7a and the injection–production working system of the corresponding period were the input of the Ctrl-CLSTM, Figure 7b shows the predicted results when the original injection–production working system was applied, and Figure 7c shows the predicted results when a different injection–production working system was applied. The first row is the numerical simulation result ($\hat{X}_t$), the second row is the Ctrl-CLSTM prediction result (X_t), the third row is the absolute value of the error ($\hat{X}_t - X_t$) between the two images and the prediction time is 360, 720, 1080, 1440, 1800 and 2160 days from left to right according to columns. When using the original injection–production working system, the maximum error is 0.14, and the error in most areas is less than 0.07. After changing the injection–production working system, the maximum error is 0.2, and the error in most areas is less than 0.1. It can be seen that with the control gate, the final state of oil saturation is changed with a high prediction accuracy when changing the injection–production parameters. Therefore, the Ctrl-CLSTM can obtain different prediction results through the control gate when changing the injection–production parameters, and the injection–production parameters can be adjusted reasonably based on this reference, which can improve the production potential of oil wells.



(**a**) Input sequence



(**b**) Predicted results

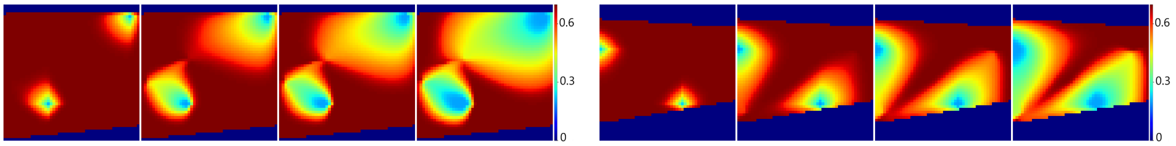


(**c**) Predicted results with changed working system

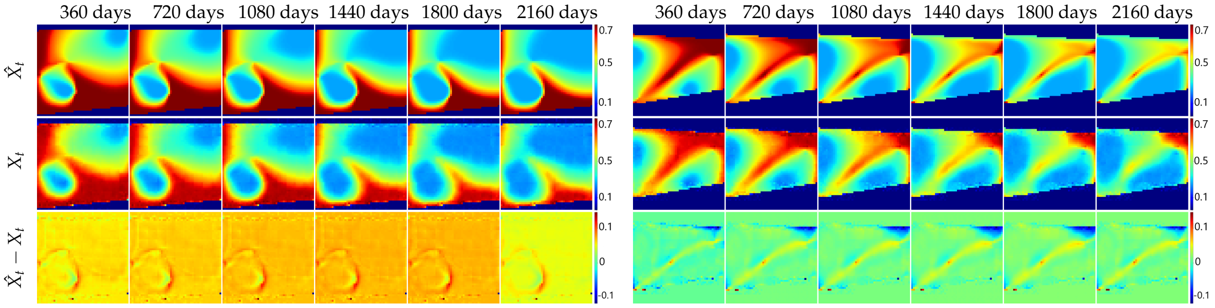
Figure 7. Prediction examples on different injection–production working systems.

5. Performance on Another Reservoir Dataset

To demonstrate the universality of the Ctrl-CLSTM, we performed experiments on another reservoir dataset generated by a numerical simulation. In this reservoir model, there were 3000 oil saturation sequences, and each sequence contained 10 images of size 50 × 50. We took the first four images as input sequence and the last six images as prediction result. In addition, these sequences had their respective injection–production parameters. In Figure 8, (a) is the input sequence of (c), (b) is the input sequence of (d); similarly, the injection–production working system in the corresponding period was also one of the inputs to the network. The rows represent the numerical simulation result ($\hat{X}_t$), the prediction result (X_t) of the Ctrl-CLSTM, and the error ($\hat{X}_t - X_t$), while the columns indicate the number of days to predict oil saturation from 360 days to 2160 days.



(**a**) Input sequence (**b**) Input sequence



(**c**) Prediction results with the input sequence (a) (**d**) Prediction results with the input sequence (b)

Figure 8. Prediction examples.

After consecutive experiments, it can be seen that the errors between both sequences are less than 0.1 in most regions, indicating that the Ctrl-CLSTM is capable of accurately modeling these different reservoirs. From Table 4, we can conclude that the value of SSIM is over 96% and the MSE is around 21. These data are very close to those derived from our experiments done on the real Tarim dataset above, proving the compatibility of the Ctrl-CLSTM with different reservoirs.

Table 4. Quantitative results on two sequences.

Input Sequence	SSIM (%)	MSE	LPIPS	PSNR (dB)
(a)	96.59	20.32	2.76	44.91
(b)	97.27	21.07	2.59	45.26

6. Conclusions

As an important application of deep learning, generating future sequence from historical continuous sequence has become a hot topic of interest. In this paper, we proposed a novel unsupervised network named Ctrl-CLSTM for spatiotemporal predictive learning tasks, which realized the unified modeling of the spatiotemporal deformations simultaneously. The core part of the Ctrl-CLSTM was the dual-memory state transmit mechanism that made temporal features propagate horizontally over time and spatial features span stacked network layers vertically. Furthermore, we innovatively introduced a control gate structure in the spatiotemporal sequence prediction network to cope with a special sequence such as oil saturation. The Ctrl-CLSTM took the existing oil saturation sequence, well location maps and injection–production parameters as input to predict the future oil saturation. When changing the injection–production parameters, the Ctrl-CLSTM can obtain different prediction results, which can save the oilfield prediction time and help to enhance the oil recovery.

In order to capture the trend of oil saturation more accurately, the Ctrl-CLSTM requires the historical data to be used in the prediction; therefore, it is mainly used in the secondary oil recovery stage to adjust the injection–production parameters continuously, but it cannot be applicable to new reservoirs that are completely undeveloped. Furthermore, when using the Ctrl-CLSTM to generate oil saturation, we did not consider sufficient factors that affect the variation of oil saturation, and in the future, we will consider more influencing factors and adapt it to more application scenarios.

Author Contributions: Conceptualization, F.L. and Y.D.; methodology, F.L.; software, F.L.; validation, Y.Z. and F.L.; formal analysis, F.L.; investigation, F.L.; resources, Q.W.; data curation, Q.W.; writing—original draft preparation, F.L.; writing—review and editing, F.L.; visualization, Y.Z.; supervision, Y.D.; project administration, Y.D.; funding acquisition, Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities (20CX05016A) and the Major Scientific and Technological Projects of CNPC under Grant ZD2019-183-007.

Data Availability Statement: The data are not publicly available for the following reasons: for the data provided by Tarim, we only use the data for the prediction of oil saturation in the Tarim project, and not for any other projects or purposes.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, Z.; Zhang, G.; Jin, Y. Quantitative characterization of oil recovery and pore volume of water injected in water flooding. *J. China Univ. Pet.* **2021**, *45*, 94–100.
2. Zhong, Z.; Sun, A.Y.; Wang, Y.; Ren, B. Predicting field production rates for waterflooding using a machine learning-based proxy model. *J. Pet. Sci. Eng.* **2020**, *194*, 107574. [CrossRef]
3. Zhang, K.; Chen, G.; Xue, X.; Zhang, L.; Sun, H.; Yao, C. A reservoir production optimization method based on principal component analysis and surrogate model. *J. China Univ. Pet.* **2020**, *44*, 90–97.
4. Yang, Y.; Yang, H.; Tao, L.; Yao, J.; Wang, W.; Zhang, K.; Luquot, L. Microscopic determination of remaining oil distribution in sandstones with different permeability scales using computed tomography scanning. *J. Energy Resour. Technol.* **2019**, *141*, 092903. [CrossRef]
5. Liu, P.; Zhang, Z.; Wang, X.; Pan, M.; Shi, Y. Revelation on remaining oil and full length hydraulic simulation of single sand bodies based on 3D models in low permeability heterogeneous reservoirs. *J. China Univ. Pet.* **2017**, *41*, 56–64.
6. Liu, Y.; Hou, S.; Wang, Z.; Yu, X.; Li, S. Reservoir prediction and application based on AVO three parameter seismic inversion under condition of dense well pattern. *Fault-Block Oil Gas Field* **2020**, *27*, 750–753.
7. Liu, T.; Li, C.; Liu, C.; Sun, Q. Experimental simulation of remaining oil distribution in combined debouch bar of delta front reservoir. *J. China Univ. Pet.* **2018**, *42*, 1–8.
8. Gu, J.; Ren, Y.; Wang, Y. Prediction methods of remaining oil plane distribution based on machine learning. *J. China Univ. Pet.* **2020**, *44*, 39–46.
9. Tang, M.; Liu, Y.; Durlofsky, L.J. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. *J. Comput. Phys.* **2020**, *413*, 109456. [CrossRef]
10. Shi, X.; Yeung, D.Y. Machine learning for spatiotemporal sequence forecasting: A survey. *arXiv* **2018**, arXiv:1808.06865.
11. Zhang, K.; Wang, Y.; Li, G.; Ma, X.; Cui, S.; Luo, Q.; Wang, J.; Yang, Y.; Yao, J. Prediction of Field Saturations Using a Fully Convolutional Network Surrogate. *SPE J.* **2021**, *26*, 1824–1836. [CrossRef]
12. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.c. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In Proceedings of the Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, Montreal, QC, Canada, 7–12 December 2015.
13. Meng, F.; Song, T.; Xu, D.; Xie, P.; Li, Y. Forecasting tropical cyclones wave height using bidirectional gated recurrent unit. *Ocean Eng.* **2021**, *234*, 108795. [CrossRef]
14. Shi, X.; Gao, Z.; Lausen, L.; Wang, H.; Yeung, D.Y.; Wong, W.k.; Woo, W.c. Deep learning for precipitation nowcasting: A benchmark and a new model. In Proceedings of the Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, Long Beach, CA, USA, 4–9 December 2017.

15. Wang, Y.; Long, M.; Wang, J.; Gao, Z.; Yu, P.S. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. In Proceedings of the Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, Long Beach, CA, USA, 4–9 December 2017.
16. Wang, Y.; Wu, H.; Zhang, J.; Gao, Z.; Wang, J.; Yu, P.S.; Long, M. PredRNN: A recurrent neural network for spatiotemporal predictive learning. *arXiv* **2021**, arXiv:2103.09504.
17. Feng, Q.; Li, Y.; Wang, S. Predicting gas migration development using deep convolutional generative adversarial network. *J. China Univ. Pet.* **2020**, *44*, 20–27.
18. Hu, Y.; O'Donncha, F.; Palmes, P.; Burke, M.; Filgueira, R.; Grant, J. A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales. *arXiv* **2021**, arXiv:2108.11875.
19. Tang, M.; Liu, Y.; Durlofsky, L.J. Deep-learning-based surrogate flow modeling and geological parameterization for data assimilation in 3D subsurface flow. *Comput. Methods Appl. Mech. Eng.* **2021**, *376*, 113636. [CrossRef]
20. Gnewuch, M.; Hebbinghaus, N. Discrepancy bounds for a class of negatively dependent random points including Latin hypercube samples. *Ann. Appl. Probab.* **2021**, *31*, 1944–1965. [CrossRef]
21. Dong, Y.; Zhang, Y.; Liu, F.; Cheng, X. Reservoir Production Prediction Model Based on a Stacked LSTM Network and Transfer Learning. *ACS Omega* **2021**, *6*, 34700–34711. [CrossRef]
22. Sukhbaatar, S.; Weston, J.; Fergus, R. End-to-end memory networks. In Proceedings of the Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, Montreal, QC, Canada, 7–12 December 2015.
23. Wu, J.; Li, Z.; Sun, Y.; Cao, X. Neural network-based prediction of remaining oil distribution and optimization of injection-production parameters. *Pet. Geol. Recovery Effic.* **2020**, *27*, 85–93.
24. Yang, X. Waterflood Development Effect Study of Heterogeneous Reservoir. *Spec. Oil Gas Reserv.* **2019**, *26*, 152–156.
25. Dağdeviren, A.; Kürüz, F. Special Real and Dual Matrices with Hadamard Product. *J. Eng. Technol. Appl. Sci.* **2021**, *6*, 127–134. [CrossRef]
26. Bock, S.; Weiß, M. A proof of local convergence for the Adam optimizer. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
27. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
28. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
29. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 586–595.