

Article

Reservoir Characterization and Productivity Forecast Based on Knowledge Interaction Neural Network

Yunqi Jiang¹, Huaqing Zhang², Kai Zhang^{1,3,*}, Jian Wang^{2,*}, Shiti Cui⁴, Jianfa Han⁴, Liming Zhang¹ and Jun Yao¹

¹ School of Petroleum Engineering, China University of Petroleum East China—Qingdao Campus, Qingdao 266580, China; jiangyunqi@s.upc.edu.cn (Y.J.); zhangliming@upc.edu.cn (L.Z.); rcogfr_upc@126.com (J.Y.)

² College of Sciences, China University of Petroleum East China—Qingdao Campus, Qingdao 266580, China; zhh@upc.edu.cn

³ School of Science, Qingdao University of Technology, Qingdao 266580, China

⁴ Exploration and Development Research Institute of PetroChina Tarim Oilfield Company, Korla 841000, China; cui shiti-tlm@petrochina.com.cn (S.C.); hanjf-tlm@petrochina.com.cn (J.H.)

* Correspondence: zhangkai@upc.edu.cn (K.Z.); wangjiannl@upc.edu.cn (J.W.)

Abstract: The reservoir characterization aims to provide the analysis and quantification of the injection-production relationship, which is the fundamental work for production management. The connectivity between injectors and producers is dominated by geological properties, especially permeability. However, the permeability parameters are very heterogenous in oil reservoirs, and expensive to collect by well logging. The commercial simulators enable to get accurate simulation but require sufficient geological properties and consume excessive computation resources. In contrast, the data-driven models (physical models and machine learning models) are developed on the observed dynamic data, such as the rate and pressure data of the injectors and producers, constructing the connectivity relationship and forecasting the productivity by a series of nonlinear mappings or the control of specific physical principles. While, due to the “black box” feature of machine learning approaches, and the constraints and assumptions of physical models, the data-driven methods often face the challenges of poor interpretability and generalizability and the limited application scopes. To solve these issues, integrating the physical principle of the waterflooding process (material balance equation) with an artificial neural network (ANN), a knowledge interaction neural network (KINN) is proposed. KINN consists of three transparent modules with explicit physical significance, and different modules are joined together via the material balance equation and work cooperatively to approximate the waterflooding process. In addition, a gate function is proposed to distinguish the dominant flowing channels from weak connecting ones by their sparsity, and thus the inter-well connectivity can be indicated directly by the model parameters. Combining the strong nonlinear mapping ability with the guidance of physical knowledge, the interpretability of KINN is fully enhanced, and the prediction accuracy on the well productivity is improved. The effectiveness of KINN is proved by comparing its performance with the canonical ANN, on the inter-well connectivity analysis and productivity forecast tasks of three synthetic reservoir experiments. Meanwhile, the robustness of KINN is revealed by the sensitivity analysis on measurement noises and wells shut-in cases.

Keywords: reservoir characterization; productivity prediction; machine learning; knowledge interaction neural network; embedded model

MSC: 37M10



Citation: Jiang, Y.; Zhang, H.; Zhang, K.; Wang, J.; Cui, S.; Han, J.; Zhang, L.; Yao, J. Reservoir Characterization and Productivity Forecast Based on Knowledge Interaction Neural Network. *Mathematics* **2022**, *10*, 1614. <https://doi.org/10.3390/math10091614>

Academic Editors: Wen Zhang, Xiaofeng Xu, Jun Wu and Kaijian He

Received: 16 March 2022

Accepted: 29 April 2022

Published: 9 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In a waterflooding reservoir, the subsurface flow is invisible and influenced by the heterogenous geophysical properties, such as the porosity, compressibility, and especially permeability. As an important content of reservoir characterization, the inter-well connectivity analysis aims to quantify the contribution from an injection well to a production well, so as to reflect the relative permeability strength of the flowing channels. Based on the analysis of inter-well connectivity, the oil field enables the adjustment of the hydrodynamics, such as water shutoff, profile control, and well pattern optimization [1–5]. Commercial simulators can predict production and characterize reservoirs accurately. However, geological information is essential for the model development by simulators, which is difficult and expensive to obtain in practice. In addition, the simulation for complex reservoirs is pretty time-consuming, usually taking several hours or even days [6–8]. For most oil fields, the injection and production rates are often available, on which the simplified reservoir simulation models can be established.

Generally, characterization approaches for the inter-well connecting relationship can be classified into three categories:

(1) Statistical and signal processing methods. These methods are based on statistical analysis and signal processing techniques. Spearman rank correlations [9] were presented to measure the relationship between injectors and producers, while the authors also pointed out that this method was not completely robust and nor were the influence factors fully understood. Tian and Horne [10] proposed a modified Pearson's correlation coefficient method to capture the influence from injectors to producers, showing a more precise inter-well characterization ability than the Spearman rank correlation method. Wavelet analysis was adopted to infer the connecting relationships [11], revealing new insights into the inter-well connectivity analysis. Some novel signal processing approaches, like cross-correlation, spectral analysis, magnitude-squared coherence, and periodogram were also applied to infer the inter-well communication [12]. Although these methods have high computational efficiency by analyzing the correlation and mapping relationships between injection and production signals, they are not established on the physical laws of waterflooding. Therefore, the robustness of these methods is hard to be guaranteed, and these models are often combined together or served as complementary methods to reduce the uncertainty [9,11,12].

(2) Machine learning methods. These methods usually quantify inter-well connectivity through model parameters. Panda and Chopra [13] proposed a related approach, using artificial neural networks (ANNs) to estimate the interactions between injectors and producers, while the geological and geostatistical data were required to determine the model parameters. Artun [14] evaluated the inter-well connectivity via the products of weight matrices in ANNs, providing a new perception of the inter-well connectivity analysis. While Jensen [15] commented that Artun's ANNs model can't reflect the physical mechanism of the waterflooding process, so it was unclear for ANN's performance on the field disturbances (e.g., temporary shut-in or completion). Even though these machine learning (ML) methods are capable of inferring the inter-well connectivity via their strong nonlinear mapping abilities, they are considered as "black box" models, for their weakness in physical interpretation, limiting their practical applications.

(3) Physical models. These models are established on the physical process and derived from corresponding physical laws. Yousef et al. [16] used the capacitance resistance model (CRM or CRMIP) to reflect the connectivity and time lag between injector and producer pairs, which derived from the material balance equation and linear productivity model. Compared with the multiple linear regression (MLR) model [17], it considered the effects of compressibility and transmissibility. Based on the work of CRM, a series of models were introduced, such as the capacitance resistance model for the producer control volume (CRMP) [18] and the capacitance resistance model for a tank or field control volume (CRMT) [19] and. Different from CRMIP, CRMT and CRMP assign each drainage volume or each tank to a constant time delay, making the production signals react synchronously

with the signals of all injectors. In addition, Sayarpour [20] proposed a CRM-blocks model, dividing the drainage volume into several blocks to calculate the flow rate. However, the complexity of the CRM-blocks model is inevitably increased, since it simulates the flowing process block by block, limiting its applications in complex cases. To infer the inter-well connectivity from multilayers, the multilayer CRM (ML-CRM) [21] was proposed, by modeling the injected fluids flowing across different layers with the help of production logging tools (PLT). Besides, Zhao et al. [22] presented an inter-well numerical simulation model (INSIM) to approximate the performance of waterflooding reservoirs. INSIM is derived from the mass material balance and front tracking equations, which consist of inter-well control units considering transmissibility and control pore volume. Moreover, considering more complex cases, such as the conversion from a producer to an injector, INSIM-FT [23] was designed; INSIM-FT-3D was proposed to simulate the flow in three dimensions with gravity [24]. Recently, INSIM-FPT [25] has been presented to reveal the inter-well connectivity via history matching data instead of the reservoir petrophysical properties. These physical models have clear physical assumptions and can be applied in other aspects, such as production optimization [26–29].

With the high computing efficiency, strong fitting ability, and excellent prediction accuracy, ML models have been widely utilized in the oil industry [28,30–36]. As an important kind of ML approach, ANNs enable to learn the complex mapping from the input variables to the desired output variables, by adjusting the weights of the internal synapses [37,38]. Nonetheless, the lack of trustworthiness is a big challenge for the further development of ANNs in real applications, since these models seldom consider physical knowledge and the model parameters do not have physical implications.

To improve the model reliability and generalizability, many researchers have tried to associate physical knowledge with ANNs in practical applications. A physics-guided neural network (PGNN) [39] was proposed to simulate the lake's temperature, using the results of physical models and leveraging physical rules to improve the scientific consistency of neural networks. In [40], physics-informed neural networks (PINN) were proposed, integrating the partial differential equation (PDE), boundary condition (BC), and initial condition (IC) into the objective function. PINN was improved in [41], which learned parameters and constitutive relationships in subsurface flow by minimizing the PDE (Darcy or Richards equation) residual. To reduce the sensitivity of initial parameters and too many iterations of primary PINN, a modified genetic algorithm is adopted in the model's optimization scheme, effectively resolving the linear elastic problems in the solid mechanics [42]. Moreover, a theory-guided neural network (TgNN) [43] was proposed to simulate the subsurface flow, which considered not only PDE, BC, and IC, but also expert knowledge and engineering controls. Csiszár et al. [44] combined continuous logic rules and multicriteria decision operators with networks, providing the semantic meaning for the values of the activation functions. In a social recommendation system, a Knowledge-aware Coupled Graph Neural Network (KCGN) is proposed, coupling the inter-dependent knowledge between items and users with the machine learning framework, which shows great performance on several real-world datasets [45]. On the one hand, most interpretable methods associate the physical knowledge with the objective function in the form of regularization terms, thus enforcing the neural networks to make predictions within certain physical constraints. However, penalty terms inevitably lead to the increase of hyperparameters, which are difficult to determine and negative to the model stability. On the other hand, the results of ML models are more consistent with the physical reality to a certain extent, yet they have a weak effect on strengthening the physical meaning of model parameters. Therefore, the "black box" can only be opened when ANNs have deeper physical interaction from the parameter level. Still, to the best of our knowledge, in terms of inter-well connectivity analysis problems, there are few studies focused on integrating the physical information with ANNs, not to mention assigning the knowledge to model parameters.

The goal of this study is to improve the accuracy and stableness of the inter-well connectivity characterization and enhance the prediction precision on well productivity, by combining the physical knowledge with machine learning techniques. The main contributions of this paper are outlined as follows.

(1) An innovative neural network is proposed to handle the reservoir characterization and productivity forecast problems, in which the material balance equation is embedded via three high transparent modules, thereby ensuring the physical sense of model parameters.

(2) A gate function is designed to evaluate the contributions from input signals to the output signals, which avoids the complex constraint optimization and guarantees the interpretability of function values.

(3) KINN reveals a successful paradigm to enhance the generalization capability and interpretability by integrating physical knowledge within the model architectures, which can be easily extended to a series of neural networks.

The rest of this paper is arranged as follows. In Section 2, we introduce the theoretical foundation (the material balance equation) of KINN. Then we provide the detailed workflow and explanation of model structures of KINN in Section 3. In Section 4, we reveal the effectiveness of KINN by comparing its performance with classical ANN on three reservoir simulation experiments and test the model's sensitivity to noisy data. Finally, we summarize this paper and get some conclusions in Section 5.

2. Methods

The material balance equation is the basic principle for a closed waterflooding reservoir, which describes the relationship between inflow, outflow, and the changes of flow among the water drainage volumes of the geological system. The inter-well connectivity analysis aims to generate a quantitative evaluation of the connecting strength for each injector-producer pair. During the waterflooding process, considering a single injector and single producer case, the material balance equation is:

$$C_t V_p \frac{d\bar{p}}{dt} = i(t) - q(t), \quad (1)$$

where C_t is the total compressibility; V_p represents the drainage pore volume; \bar{p} is the average pressure of V_p ; t represents the timestep; $i(t)$ and $q(t)$ are the vectors denote the injection rate and production rate, respectively. Equation (1) assumes that the total compressibility is a small constant, and no fluids flow into or out of the drainage volume. Assume that there is a case with M injectors and N producers, using superposition in space of M injectors and ignoring the response of the production signals before injection, the production rate for producer j is given by:

$$\sum_{k=1}^M C_{tkj} V_{pkj} \frac{d\bar{p}_{kj}}{dt} = \sum_{k=1}^M \lambda_{kj} i_k(t) - q_j(t), \quad (2)$$

where k is the injector index, $k = 1, 2, \dots, M$; j is the producer index, $j = 1, 2, \dots, N$; and λ_{kj} denotes the connectivity value between injector k and producer j .

3. Knowledge Interaction Neural Network (KINN)

As shown in Figure 1, KINN is a first-principle-based model with modularized architectures, where each module keeps a one-to-one correspondence with each item of the material balance equation. According to Equation (2), KINN is established on each producer, considering the flow from all injectors and the influence caused by the compressibility of the control volume. There are two input modules in KINN, named injection regulator module (IRM) and control volume module (CVM), respectively. IRM corresponds to the injection item in the material balance equation, using a gate function layer to quantitatively measure the contribution from each injector to the analyzed producer. CVM is used for approximating the fluid change rate in the control volume via a series of fully connected layers. Then, the model output of the analyzed producer is controlled by the output system,

called the production monitor module (PMM). It aims to calculate the estimated production rate via the outputs of IRM and CVM according to the material balance equation. In brief, within the framework of the material balance equation, the three modules of KINN interact with physical knowledge corporately, then simulate the water flooding process and characterize inter-well connectivity through network parameters (gate functions).

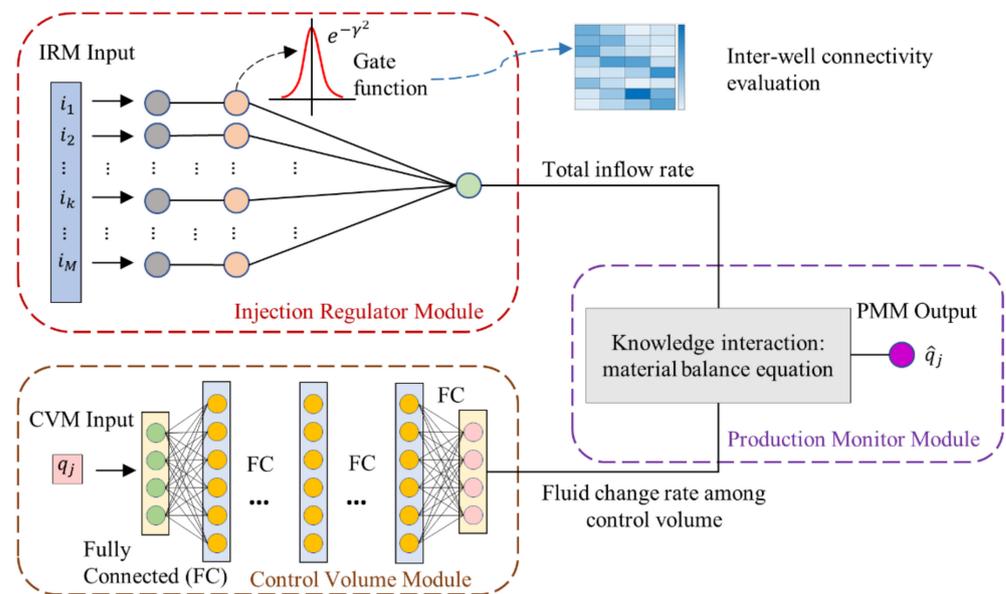


Figure 1. The architecture of KINN.

3.1. Injection Regulator Module (IRM)

The injection regulator module is an essential part of the input system in KINN, responsible for measuring the total flow from all injectors to the target producer (the first item on the right-hand side of Equation (2)) and inferring the inter-well connectivity by gate functions.

Let's assume that $I = [i_1, i_2, \dots, i_k, \dots, i_M]^T$ is the well water injection rate (WIR) data of M injectors, and $Q = [q_1, q_2, \dots, q_j, \dots, q_N]^T$ is the well liquid production rate (LPR) data of N producers, where i_k and q_j are vectors. As shown in Figure 1, the input data of IRM is I , followed by a gate function layer defined as:

$$g(\gamma_{kj}) = e^{-\gamma_{kj}^2}, \tag{3}$$

$$g(\gamma)_{M \times N} = \begin{pmatrix} g(\gamma_{11}) & \cdots & g(\gamma_{1N}) \\ \vdots & g(\gamma_{kj}) & \vdots \\ g(\gamma_{M1}) & \cdots & g(\gamma_{MN}) \end{pmatrix}, \tag{4}$$

where $g(\gamma)_{M \times N}$ denotes the inter-well connectivity matrix; $g(\gamma_{kj})$ is the connectivity value between injector k and producer j ; and γ_{kj} denotes the independent variable of $g(\gamma_{kj})$.

The output of IRM is calculated by:

$$I_j = \sum_{k=1}^M g(\gamma_{kj}) \cdot i_k, \tag{5}$$

where \cdot represents the product between one scalar and one vector and I_j denotes the comprehensive injection rate for producer j .

The multiplicity of the solution is a big challenge for inter-well connectivity analysis since it is a typical inverse problem. To reduce the multiplicity caused by the initialization of model parameters, Pearson correlation is employed in the initialization of $\gamma_{M \times N}$:

$$\rho(I, Q) = \frac{cov(I, Q)}{\sigma_I \cdot \sigma_Q}, \tag{6}$$

where $\rho(I, Q)$ and $cov(I, Q)$ are the correlation matrix and the covariance matrix between I and Q ; σ_I, σ_Q are the standard deviations of I and Q . By calculating the reciprocal of each element in $\rho(I, Q)$, we can get $\gamma_{M \times N}$:

$$\gamma_{M \times N} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1N} \\ \vdots & \gamma_{kj} & \vdots \\ \gamma_{M1} & \cdots & \gamma_{MN} \end{pmatrix}, \tag{7}$$

where γ_{kj} is the reciprocal of the Pearson correlation coefficient between i_k and q_j .

For producer j , its initialized independent variable of the gate function, $\gamma_{M \times 1}$, is shown in Figure 2a, which is column j of $\gamma_{M \times N}$. During the initialization process, if the relationship between the signals of injector k and producer j is strong, their correlation coefficient would be big, and γ_{kj} would be close to 0, which means its gate value, $g(\gamma_{kj})$, would be large. In this way, each good pair would be assigned a fixed connectivity value according to the relation strength, which helps to decrease the multiplicity of the inter-well connectivity.

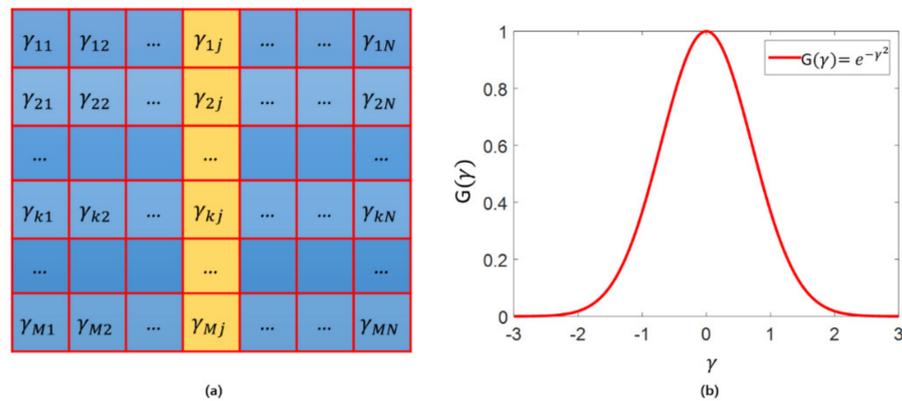


Figure 2. The curve of the gate function and its independent variables: (a) independent variables of the gate function; (b) curve of the gate function.

There are two major purposes for the presentation of the gate function.

(1) One purpose is that the negative connectivity value can be eliminated by the gate function, as its value is strictly constrained in $(0, 1]$, as shown in Figure 2b. Therefore, the physical meaning of the gate function values can be guaranteed (the contribution from injector to producer). Moreover, γ_{kj} can be updated by normal unconstrained optimization methods, ensuring the fast convergence speed of KINN.

(2) The other purpose is that the proposed gate function has great significance for the stableness of KINN. As can be seen in Figure 2b, the gate function has a good sparsity, thereby only the strong connecting well pairs would be assigned to big connectivity values and the values of the weak or none connecting pairs would be maintained at a low level. In the machine learning field, it is common sense that the sparsity feature is helpful to support the model’s robustness.

In IRM, for every producer, only $\gamma_{M \times 1}$ requires optimization, and the gate function layer can be used as inter-well connectivity indicators directly, once KINN has finished training.

3.2. Control Volume Module (CVM)

As shown in Figure 1, the control volume module is another part of the input system, aiming at computing the flow among the control volume, the left-hand side of Equation (2). In reservoir waterflood simulation, if the average reservoir pressure, \bar{p} , is a constant, the linear productivity model is often used to describe the relationship between production rate (q) and bottom hole pressure (p_{wf}). However, in an unstable flow case, q changes continuously, and the linear prediction model cannot represent the exact mapping relationship between q and p_{wf} . Figure 3 illustrates two inflow performance relationship (IPR) curves: one denotes the actual curve in the unstable flow case, and another represents the curve of the linear productivity model. Obviously, the linear model cannot provide a precise map between q and p_{wf} in the unstable flow case. To overcome this defect, some fully connected layers (some layers use nonlinear activation functions) are utilized in CVM to learn the nonlinear mapping relationships between q and p_{wf} . The motivation behind CVM is that the input variables can be mapped into a nonlinear space via the activation functions of nonlinear layers, thereby combining the connected weights to approximate the target output.

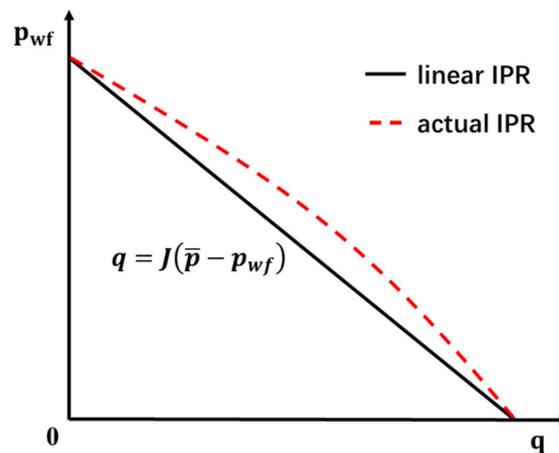


Figure 3. The inflow performance relationship (IPR) curves. The black line is the linear IPR prediction curve, and the red dashed line is the IPR curve of unstable flow.

Here, considering a single producer, we use:

$$q = \tilde{N}(\bar{p} - p_{wf}), \tag{8}$$

to represent the nonlinear relationship, where \tilde{N} denotes a nonlinear mapping. In oil fields, \bar{p} is usually unavailable, so we indicate the relationship between q and $\bar{p} - p_{wf}$ by:

$$\bar{p} - p_{wf} = N(q), \tag{9}$$

where N denotes another nonlinear mapping different from \tilde{N} . The differential form of Equation (9) with respect to the time step, t , is as follows:

$$\frac{d\bar{p}}{dt} - \frac{dp_{wf}}{dt} = \frac{dN(q)}{dt}. \tag{10}$$

Considering the case with constant BHP, Equation (10) can be simplified as:

$$\frac{d\bar{p}}{dt} = \frac{dN(q)}{dt}. \tag{11}$$

Multiply the left and right sides by $C_t V_p$ at the same time, and the change rate of the flow in the control volume is given by:

$$C_t V_p \frac{d\bar{p}}{dt} = C_t V_p \frac{dN(q)}{dt}. \tag{12}$$

For the analyzed producer j , considering M injectors, we utilize an amount of fully connected layers to approximate:

$$\sum_{k=1}^M C_{tkj} V_{pkj} \frac{d\bar{p}_{kj}}{dt} \sim Net(q_j), \tag{13}$$

where Net represents the connected layers in the network. In this paper, tansig function and Gaussian kernel function are employed as activation functions in CVM, respectively.

To sum up, several fully connected layers are used in CVM to approximate the sophisticated mapping from the LPR of the producer to the flow change rate among the control volume. Consequently, the robustness of KINN can be guaranteed even in the unstable flow case.

3.3. Production Monitor Module (PMM)

The production monitor module (PMM) is employed to calculate the liquid production rate of producer j , as shown in Figure 1. Based on the outputs generated by IRM and CVM, according to the material balance equation, the output of PMM can be given as:

$$\hat{q}_j = \Gamma_j - Net(q_j), \tag{14}$$

where \hat{q}_j denotes the estimated production rate of producer j , the second item on the right-hand side of Equation (2). In this paper, the mean square error (MSE) function is used as the loss function:

$$MSE(q_j, \hat{q}_j) = \frac{1}{T} \sum_{t=1}^T (q_j(t) - \hat{q}_j(t))^2, \tag{15}$$

where t is the time step and T denotes the number of total time steps. Therefore, the difference between model output and target output can be monitored by PMM, and the waterflood simulation can be achieved by minimizing Equation (15).

As demonstrated in Figure 1, IRM and CVM are united with PMM under the control of the material balance equation, to approximate the influence caused by water injection and compressibility, respectively. Hence, KINN enables different modules to interact physical knowledge with each other during the learning process. Additionally, the transparency of KINN is significantly improved from the underlying parameter level, and both robustness and computation efficiency are successfully combined by integrating physical information within the ML framework.

3.4. Reservoir Characterization and Productivity Prediction

For producer j , considering the effect of all injection wells, KINN can be established on I and q_j , and each injector would obtain a gate function to evaluate its connectivity value with the producer j . The workflow of the KINN training procedure is shown in Figure 4, and the pseudocode is demonstrated in Algorithm 1. Firstly, $\gamma_{M \times N}$ must be initialized via Pearson correlation method with given I and Q , and the connecting weights in CVM also need an initialization. Afterward, guided by the material balance equation, IRM and CVM would cooperatively simulate the influence caused by water injection and compressibility in the waterflooding process. In the IRM part, the input comes from the water injection rate data of all injectors, and each vector would multiply a gate function, which is the inter-well connectivity indicator and has to be optimized during the training procedure. The output of IRM is the total inflow rate, the sum of all multiplied vectors. In the CVM part, a number of fully connected layers are utilized to realize the map from liquid production rate to the fluid change rate of the control volume, whose connecting weight

matrices need optimization in the learning process. Afterward, both IRM and CVM would be combined by the material balance equation to calculate the model liquid production rate for producer j , and the loss between model output and target output can be measured by the loss function. It must be noted that the model loss comes from the outputs of both IRM and CVM, hence their physical knowledge would interact cooperatively with the model training process. During the optimization process, all the model parameters would be updated at the same time. When the stop criterion is satisfied, KINN would stop training, and the inter-well connectivity values can be inferred directly by gate functions.

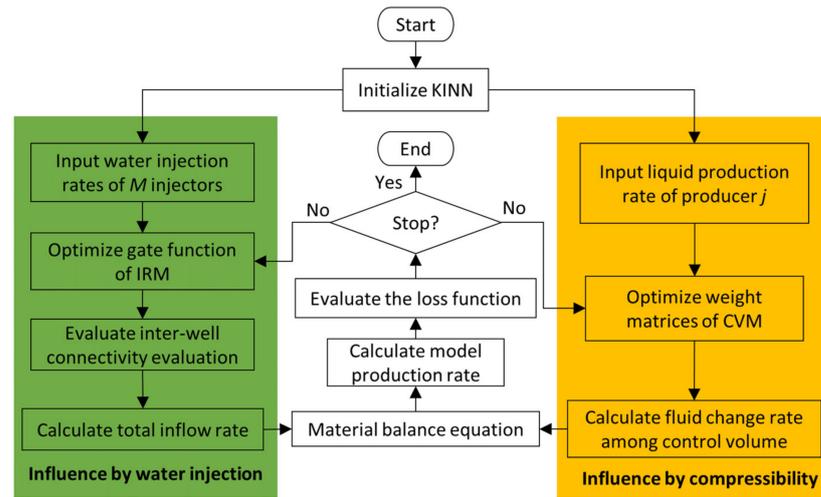


Figure 4. Flow diagram of KINN training.

Productivity forecast is the testing process of KINN, revealing its generalization performance. Unlike regular machine learning models, the liquid production rate is used as both input and output during the training procedure, as shown in Figure 1. When the training process is finished, all parameters of KINN are constants, then it can forecast the production rate \hat{q}_j with given I , by solving the nonlinear equation:

$$\hat{q}_j - \Gamma_j + Net(\hat{q}_j) = 0. \tag{16}$$

Algorithm 1: Knowledge Interaction Neural Network (KINN)

Input: I , WWIR for M injectors, and Q , WLPR for N producers

Output: \hat{q}_j

- ```

/ *** start KINN training *** /
1 Initialization $\lambda_{M \times N}$: Compute $\gamma_{M \times N}$ using database I and Q by Equations (6) and (7), and initialize the parameters of ANN in CVM
2 For $j = 1$ to N do
3 While convergence tolerance is not met
4 / *** IRM calculation *** /
4 Select the j_{th} column, $\gamma_{M \times 1}$, in $\gamma_{M \times N}$ as the independent variable of gate function;
5 Calculate the output of IRM, Γ_j , with $\gamma_{M \times 1}$ and I , using Equation (5)
6 / *** CVM calculation *** /
6 Calculate the output of CVM, $Net(q_j)$, with q_j , using Equation (13)
7 / *** PMM calculation *** /
7 Calculate the output of PMM, \hat{q}_j , using Equation (14)
8 / *** parameters update *** /
8 Evaluate the loss function using Equation (15)
9 Update $\gamma_{M \times 1}$ and weight matrices of CVM via gradient descent algorithm
10 End While
11 End For
/ *** end KINN training *** /

```
-

#### 4. Results

Three reservoir cases with various inter-well connecting conditions are studied in this section, including the streak reservoir case, the braided river reservoir case and Egg reservoir case, developed on ECLIPSE (Schlumberger Ltd., Houston, TX, USA). KINN has taken two activation functions (tansig function and Gaussian kernel function) in CVM, named KINN-tansig and KINN-Gaussian, respectively. To compare the performance of the proposed models and the classical neural network without the guidance of physical information, we demonstrate the results obtained by the single-hidden-layer feedforward neural network (SLFNN). The numbers of the input nodes and output nodes are equal to the numbers of injectors and producers. To keep the fairness of comparison, the number of hidden nodes, the activation function, the learning rate, the convergence error, and the optimization method are the same as those of KINN-tansig. The connectivity matrix of SLFNN is the normalized absolute value of the product between the input-hidden-layer weights and the hidden-output-layer weights. All three models use an Intel(R) Core (TM) i7-9700 CPU. The data are separated into two parts, where the former 80% are used in history matching and the latter 20% are utilized in productivity prediction. Table 1 shows the hyperparameters of KINN-tansig and KINN-Gaussian in three cases. To demonstrate the inter-well connectivity clearly and intuitively, the connectivity characterization results are visualized by heatmaps, where the deeper the block color, the stronger the injector-producer connectivity.

**Table 1.** The hyperparameters for KINN-tansig and KINN-Gaussian in three cases.

| Hyperparameter                           | KINN-Tansig             | KINN-Gaussian            |
|------------------------------------------|-------------------------|--------------------------|
| Learning rate                            | 0.05                    | 0.05                     |
| Number of hidden layers in CVM           | 3                       | 3                        |
| Number of neurons of each layer in CVM   | [1, 10, 1]              | [1, 10, 1]               |
| Activation function in CVM               | tansig function         | Gaussian kernel function |
| Initialization range of weights in CVM   | [0, 0.25]               | [0, 0.25]                |
| Initialization method of $\gamma$ in IRM | Pearson Correlation     | Pearson Correlation      |
| Optimization algorithm                   | Gradient descent method | Gradient descent method  |
| Convergence error (MSE)                  | $10^{-6}$               | $10^{-6}$                |

##### 4.1. The Streak Reservoir Case

This model is reconstructed from the work of [19], which consists of  $31 \times 31$  single-layer grids in the X-Y plane, with 80, 80, and 12 ft in the X, Y, and Z axes, respectively. There are 5 injectors, named I1, I2, I3, I4, I5, and 4 producers, named P1, P2, P3, P4. As shown in Figure 5, the permeability of the matrix is 5 md, except for two high-permeability streaks. One streak is between I1 and P1 of 1000 md, and the other is 500 md between I3 and P4. The normal properties of the streak model are shown in Table 2. The simulated production lasts around 1800 days, and the timestep is 5 days. Because the permeability values of I1-P1 and I3-P4 are much higher than other well pairs, the fluids are less likely to flow into P2 and P3, thus their production rates are quite low. Here, only the history matching and production prediction results of P1 and P4 are given.

**Table 2.** Properties of the streak reservoir model.

| Properties               | Value                                 |
|--------------------------|---------------------------------------|
| Model Size               | $31 \times 31 \times 1$               |
| Depth                    | 2000 m                                |
| Initial pressure         | 2000 psi                              |
| Porosity                 | 0.18                                  |
| Initial water saturation | 0.3                                   |
| Density of oil           | $900 \text{ kg/m}^3$                  |
| Viscosity of oil         | 2.0 cp                                |
| Oil compressibility      | $5.0 \times 10^{-6} \text{ bar}^{-1}$ |

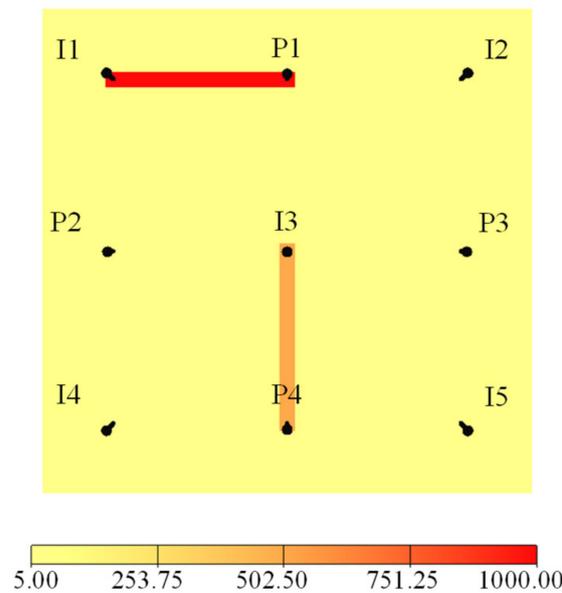
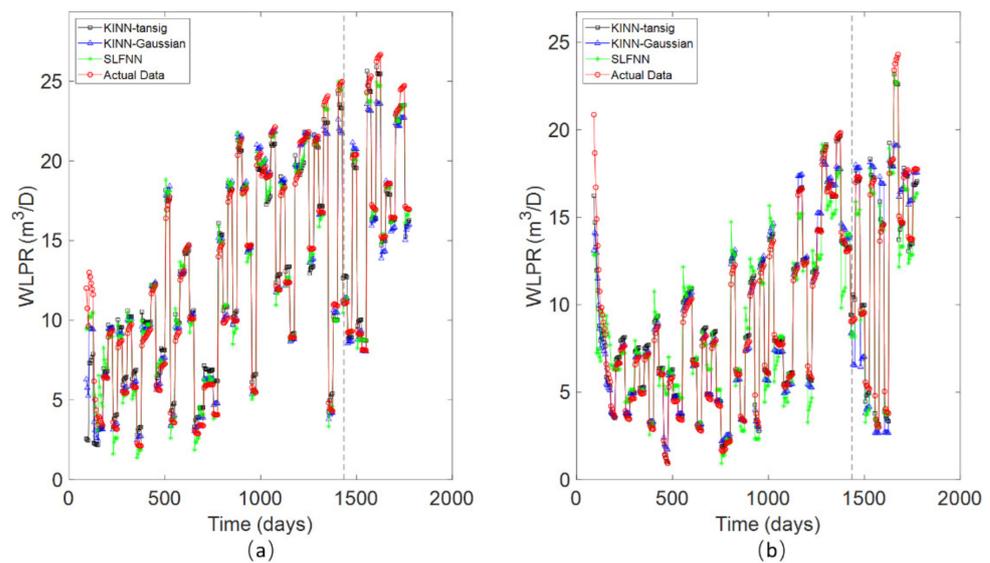


Figure 5. Permeability field of the streak reservoir case.

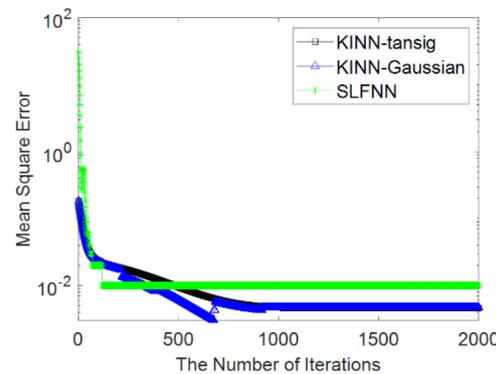
Figure 6a,b illustrate the history matching and productivity prediction results of P1 and P4, respectively. Obviously, the proposed two methods outperform SLFNN on both history matching and prediction periods. As shown in the figures, both KINN-tansig and KINN-Gaussian can obtain high fitness accuracy on P1 and P4 in the history matching period. And the two proposed models show different performances during the productivity prediction period. As illustrated in Figure 6a, there are three peaks in the prediction period of P1, which can be estimated accurately by KINN-tansig, while the forecast by KINN-Gaussian is lower than the actual values. Similar results can be found in Figure 6b, where KINN-tansig outperforms KINN-Gaussian in the prediction of P4. Figure 7 illustrates the training error (MSE) curves of KINN-tansig, KINN-Gaussian, and SLFNN, where both two proposed methods show smaller errors than SLFNN. Note that there are some fluctuations on the error curve of KINN-Gaussian, while the accuracy of KINN-Gaussian is equivalent to that of KINN-tansig when the training is finished. As shown in Table 3, KINN-tansig costs less computation time (0.3702 s) than KINN-Gaussian (2.3393 s) in this case, and the computation time of SLFNN (1.2737 s) is in the middle of the three methods. Both the history matching and prediction errors of KINN-tansig and KINN-Gaussian are around one order magnitude lower than those of SLFNN. Moreover, even though the history matching errors of KINN-tansig (0.0046) and KINN-Gaussian (0.0047) are very close to each other, the former outperforms the latter in the production prediction, with the errors of 0.0223 and 0.0256, respectively.

Table 3. The time consumption, training error (MSE) and testing error (MSE) of KINN-tansig, KINN-Gaussian and SLFNN in the streak reservoir case.

| Methods                                    | KINN-Tansig | KINN-Gaussian | SLFNN    |
|--------------------------------------------|-------------|---------------|----------|
| Computation time (training and testing)    | 0.3702 s    | 2.3393 s      | 1.2737 s |
| Error of history matching (training error) | 0.0046      | 0.0047        | 0.0976   |
| Error of prediction (testing error)        | 0.0223      | 0.0256        | 0.1832   |

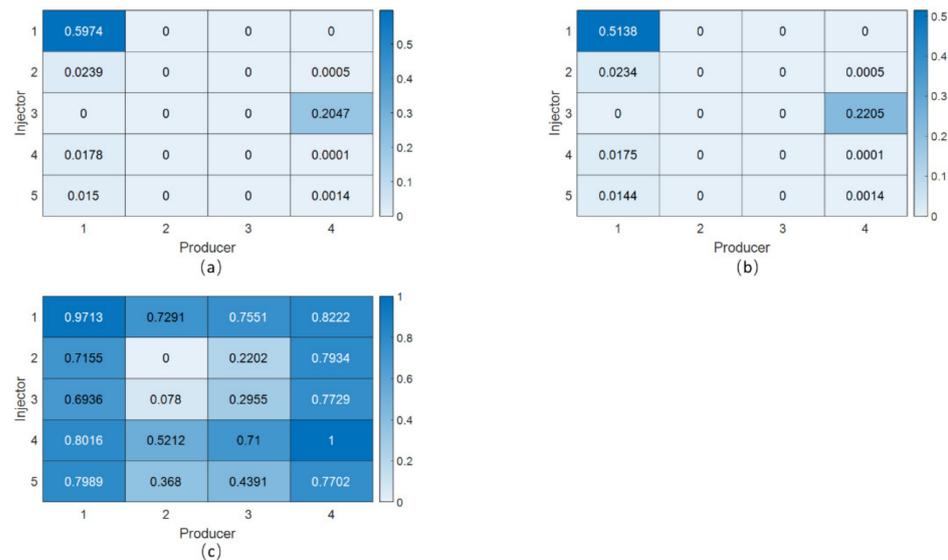


**Figure 6.** History matching results of the streak reservoir case by KINN-tansig, KINN-Gaussian and SLFNN. The black line with squares is the result obtained by KINN-tansig method; the blue line with triangles represents the result gotten by KINN-Gaussian method; the green line with stars is the results obtained by SLFNN; the red line with circles is the liquid production rate gotten by ECLIPSE; the grey vertical dashed line makes a separation of history matching period and the productivity forecast period: (a) P1; (b) P4.



**Figure 7.** The training error (MSE) curves of KINN-tansig, KINN-Gaussian and SLFNN for the streak reservoir case. The black line with squares is the error curve of KINN-tansig; the blue line with triangles represents the error curve of KINN-Gaussian; and the green line with stars represent the error curve of SLFNN.

Figure 8 illustrates the inter-well connectivity analysis results produced by three models. Undoubtedly, I1-P1 and I3-P4 should be the top and the second highest connecting well pairs according to the permeability distribution in Figure 5, which are indicated truthfully by KINN-tansig and KINN-Gaussian. In detail, the permeability of the streak of I1-P1 is 1000 md, twice larger than the streak of I3-P4, so that their corresponding connectivity values should also reflect this difference. As shown in Figure 8a, the connectivity values of I1-P1 and I3-P4 obtained by KINN-tansig are 0.5974 and 0.2047, respectively. Similarly, as illustrated in Figure 8b, KINN-Gaussian assigns I1-P1 and I3-P4 with the values of 0.5138 and 0.2205, separately. However, as demonstrated in Figure 8c, the top two connectivity values are assigned to I4-P4 (1.000) and I1-P1 (0.9713), and the value of I3-P4 (0.7729) only ranks seventh, which means SLFNN mistakenly identifies the weak connecting well pairs as the strong connecting ones. In contrast, KINN-tansig and KINN-Gaussian successfully allocate the well pairs on the low permeability area with quite small connectivity values, showing great accordance with their actual geological properties.



**Figure 8.** The heatmaps of the inter-well connectivity analysis by three models for the streak reservoir case: (a) KINN-tansig; (b) KINN-Gaussian; (c) SLFNN.

4.2. The Braided River Reservoir Case

To test the performance of KINN in other more complex cases, we have designed the braided river reservoir case, which is a classical fluvial deposition distributing in continental facies basin. There are  $100 \times 100$  single-layer grids in the braided river reservoir model, where each grid is 80, 80 and 12 ft in the X, Y and Z axes, respectively. Except for the model size, the other properties of the braided river reservoir model are the same as shown in Table 4. As shown in Figure 9, the permeability distributions are significantly different between river channels and other areas, whose permeability values are set to be 1000 md and 50 md, respectively. The simulated production lasts around 1800 days and the time step is 1 day. In this case, there are also 5 injectors, named I1, I2, I3, I4 and I5, and 4 producers, called P1, P2, P3 and P4. I1 is located on the top left corner, connecting P1 through the river channel. P2, P3 and P4 are connected with I5 by three tributaries, respectively, where the tributary between I5 and P2 is widest. Besides, the tributaries of I5-P3 and I5-P4 are of similar width, while the distance of I5-P4 is longer than that of I5-P3.

**Table 4.** The time consumption, training error (MSE) and testing error (MSE) of KINN-tansig, KINN-Gaussian and SLFNN in the braided river reservoir case.

| Methods                                    | KINN-Tansig | KINN-Gaussian | SLFNN    |
|--------------------------------------------|-------------|---------------|----------|
| Computation time (training and testing)    | 0.7417 s    | 3.4679 s      | 2.4602 s |
| Error of history matching (training error) | 0.0052      | 0.0058        | 0.0104   |
| Error of prediction (testing error)        | 0.0071      | 0.0065        | 0.0142   |

In the braided river reservoir case, even the production rates change significantly, the two proposed models are capable of matching the history of 4 producers with certain accuracy in general, as shown in Figure 10. In contrast, SLFNN shows a poorer performance than KINN-tansig and KINN-Gaussian, on both history matching and forecast tasks of all producers, where the green lines with stars often deviate from the actually observed curves (the red lines with circles). When it comes to the details, KINN-tansig and KINN-Gaussian show different performances on different producers. For instance, as can be seen in Figure 10a–d, the history matching results on P1 gotten by both two models are not as good as those on other producers, especially for KINN-tansig, whose estimated curve may be above or below the actual values. In the productivity prediction period, both KINN-tansig and KINN-Gaussian conduct results with significant fluctuations on P4, while they still show great performance on the other producers.

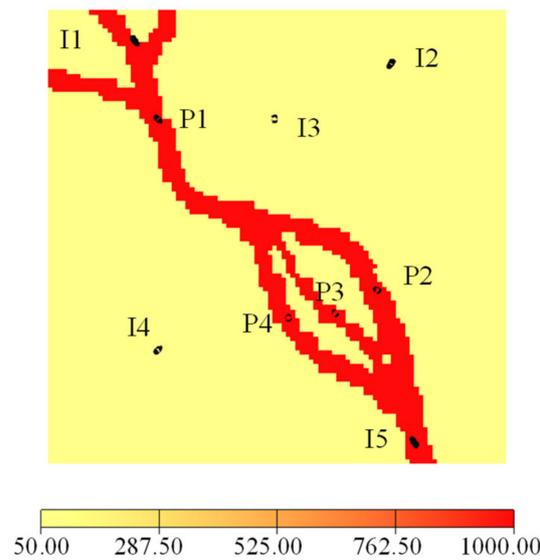


Figure 9. Permeability field of the braided river reservoir case.

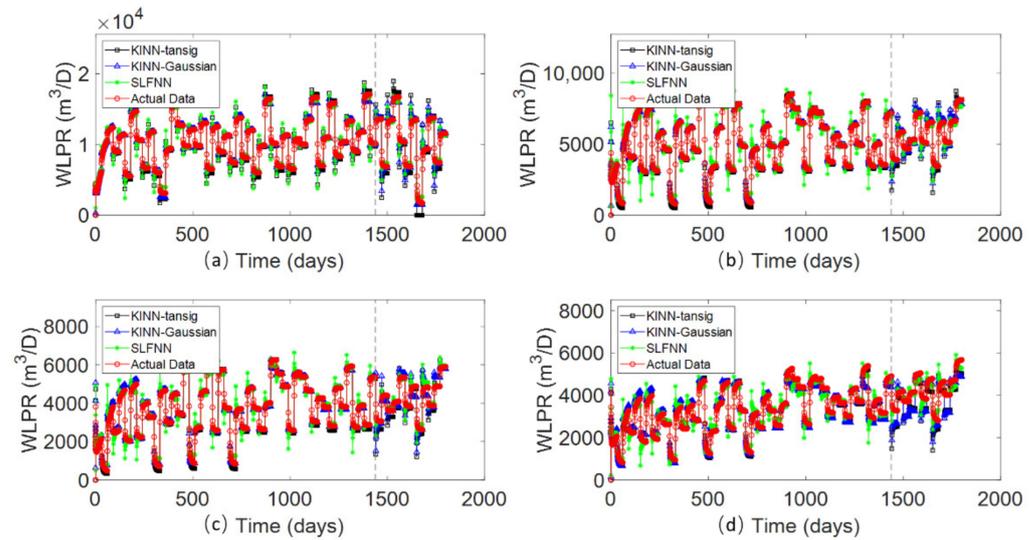
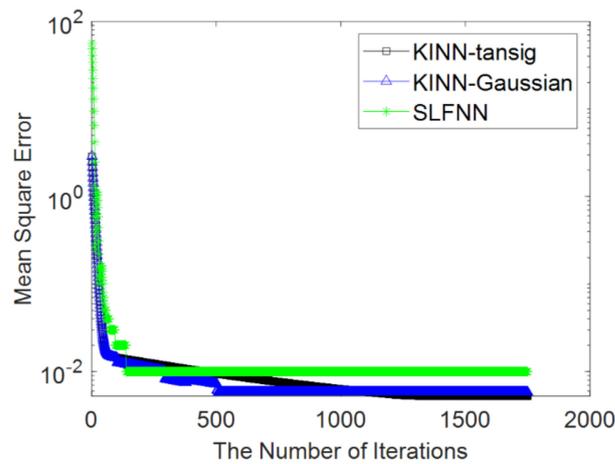


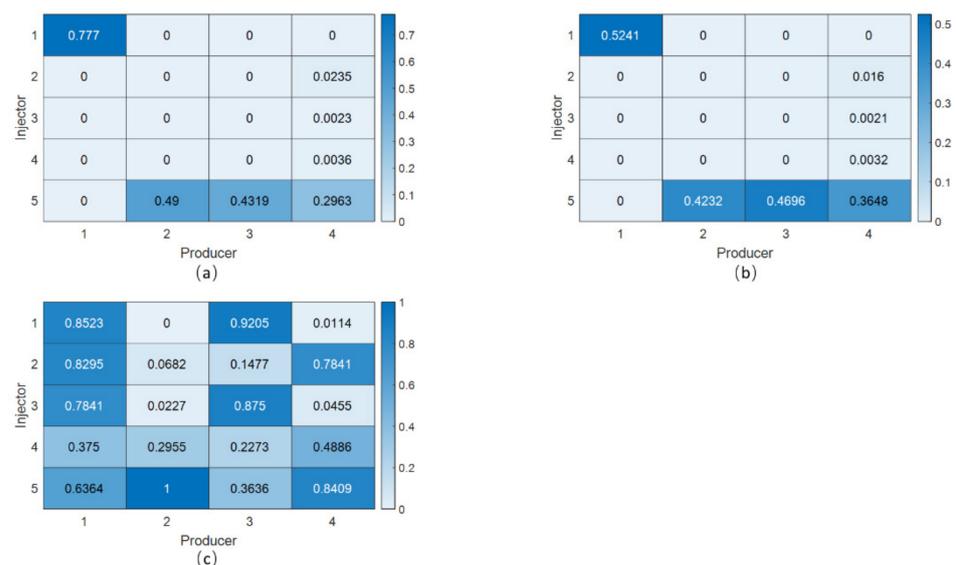
Figure 10. History matching results of the braided river reservoir case by KINN-tansig, KINN-Gaussian and SLFNN. The black line with squares is the result obtained by KINN-tansig method; the blue line with triangles represents the result gotten by the KINN-Gaussian method; the green line with stars is the results obtained by SLFNN; the red line with circles is the liquid production rate gotten by ECLIPSE; the grey vertical dashed line makes a separation of history matching period and the productivity forecast period: (a) P1; (b) P2; (c) P3; (d) P4.

As can be seen in Figure 11, KINN-tansig and KINN-Gaussian keep high computation efficiency in the braided river reservoir case, and their training errors converge to a neighborhood between 0.0050 and 0.0060, which is about a half of the error of SLFNN (0.0104). In addition, as shown in Table 4, it costs 0.7417 s for KINN-tansig and 3.4679 s for KINN-Gaussian to finish training and 2.4602 s for SLFNN. The proposed two approaches also have smaller training and testing errors than SLFNN in the braided river case.



**Figure 11.** The training error (MSE) curves of KINN-tansig, KINN-Gaussian and SLFNN for the braided river reservoir case. The black line with squares is the error curve of KINN-tansig; the blue line with triangles represents the error curve of KINN-Gaussian; and the green line with stars represent the error curve of SLFNN.

According to the permeability distribution of the braided river reservoir case, the injector-producer pairs with high flow channels are I1-P1, I5-P2, I5-P3 and I5-P4, ranking by the strength of their connecting conditions. As shown in Figure 12, these strong connecting well pairs can be revealed directly through deep color grids in heatmaps. Figure 12a demonstrates that the inter-well connectivity can get a good reflection by KINN-tansig, as the top four connectivity values and the top four high connecting well pairs are one-to-one matched, where I1-P1 is biggest with 0.777, following I5-P2, I5-P3 and I5-P4, with 0.49, 0.4319 and 0.2963, respectively. Figure 12b shows that KINN-Gaussian generates similar characterization results with KINN-tansig, except that the value of I5-P3 (0.4696) is bigger than that of I5-P2 (0.4232). SLFNN enables to characterize three strong connecting well pairs, I5-P2, I1-P1 and I5-P4, with the values of 1, 0.8523 and 0.8409. Meanwhile, I5-P3 only obtains 0.3636, which is much lower than some actually weak connecting well pairs, such as I1-P3 (0.9205), I2-P1 (0.8295), I2-P4(0.7841), I3-P1 (0.7841) and I3-P3 (0.875).



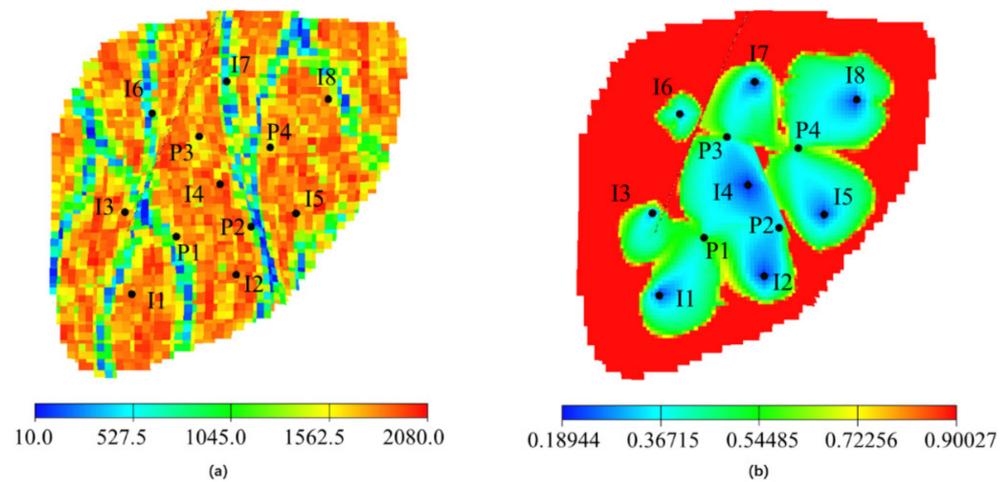
**Figure 12.** The heatmaps of the inter-well connectivity analysis by KINN-tansig and KINN-Gaussian for the streak reservoir case: (a) KINN-tansig; (b) KINN-Gaussian; (c) SLFNN.

### 4.3. Egg Reservoir Case

The initial Egg model can be seen in the work of [46], and some modifications are taken to make it more suitable for the inter-well connectivity analysis. This synthetic reservoir model consists of active 6910 grids, and the size of each grid in the X, Y, and Z directions is 8 m, 8 m, and 4 m, respectively. The important properties of Egg reservoir model are presented in Table 5. The simulated production lasts around 1200 days and the time step is 10 days. As shown in Figure 13a, there are 8 injectors and 4 producers in this case, and there are two faults in the Egg reservoir model, blocking the flow of underground fluid. In this way, the relationships between injectors and producers located on the different sides of the fault should be pretty weak. To understand the communications between injectors and producers in detail, the oil saturation distribution is demonstrated in Figure 13b.

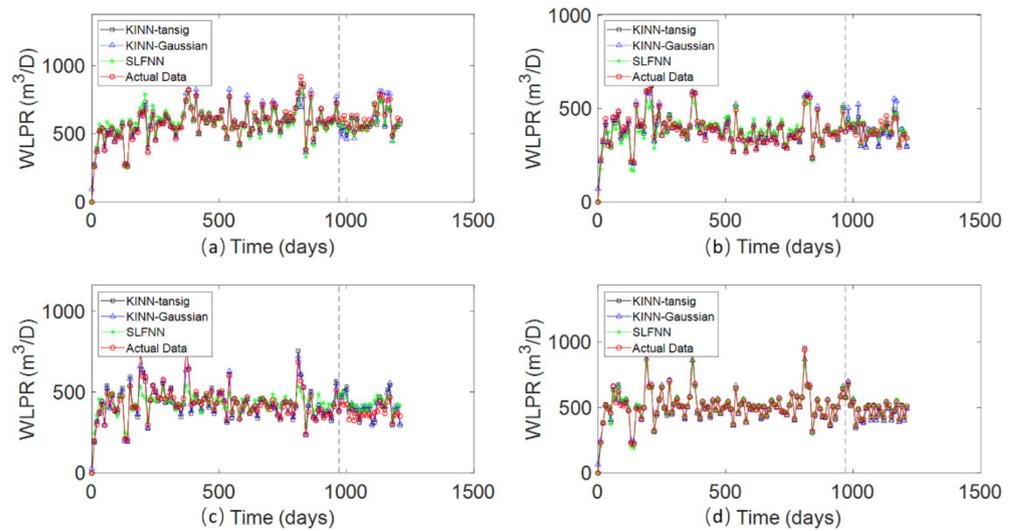
**Table 5.** Properties of Egg reservoir model.

| Properties               | Value                                    |
|--------------------------|------------------------------------------|
| Model Size               | 100 × 99 × 1                             |
| Depth                    | 4000 m                                   |
| Initial pressure         | 5765 psi                                 |
| Porosity                 | 0.2                                      |
| Initial water saturation | 0.1                                      |
| Density of oil           | 900 kg/m <sup>3</sup>                    |
| Viscosity of oil         | 2.0 cp                                   |
| Oil compressibility      | 1.0 × 10 <sup>-5</sup> bar <sup>-1</sup> |



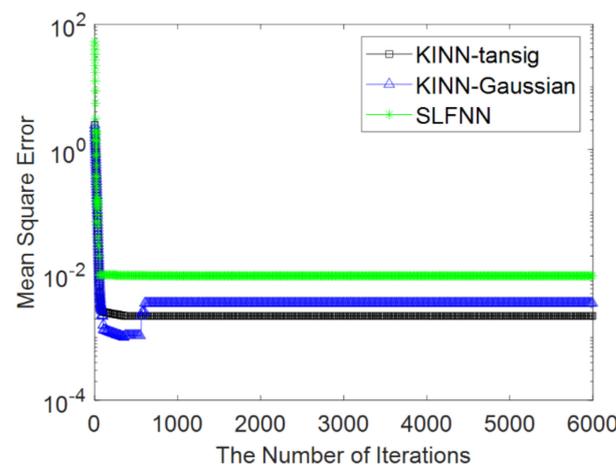
**Figure 13.** The permeability and oil saturation distribution of Egg reservoir case: (a) Permeability distribution; (b) oil saturation distribution.

As can be seen in Figure 14a–c, there are a certain number of points obtained by SLFNN (denoted by the green lines with stars) deviated from the actual ones (denoted by the red lines with circles), especially for the results on P4. Meanwhile, the proposed two methods (black lines with squares for KINN-tansig, and blue lines with triangles for KINN-Gaussian) show much better performance than SLFNN on the history matching and prediction for the LPR data of P1, P2 and P3. Besides, all three models show pretty good performance on P4, as shown in Figure 14d.



**Figure 14.** History matching results of the Egg reservoir case by KINN-tansig, KINN-Gaussian and SLFNN. The black line with squares is the result obtained by KINN-tansig method; the blue line with triangles represents the result gotten by KINN-Gaussian method; the green line with stars is the results obtained by SLFNN; the red line with circles is the liquid production rate gotten by ECLIPSE; the grey vertical dashed line makes a separation of history matching period and the productivity forecast period: (a) P1; (b) P2; (c) P3; (d) P4.

Figure 15 shows that KINN-tansig and KINN-Gaussian converge fast in the training process, where the MSE errors of both methods are reduced to less than  $10^{-2}$  within 200 iterations. Meanwhile, the initial MSE for SLFNN is much bigger than that of KINN-tansig and KINN-Gaussian, and so is the converged error. As illustrated in Table 6, KINN-tansig, KINN-Gaussian, and SLFNN all demonstrate significant computation efficiency, taking 1.1282 s, 0.8539 s, and 0.3361 s, to finish training, respectively. As expected, both KINN-tansig and KINN-Gaussian are capable of producing more accurate results in history matching (0.0022 and 0.0035) and productivity prediction (0.0171 and 0.02263, respectively) than those obtained by SLFNN (0.0097 and 0.0426).

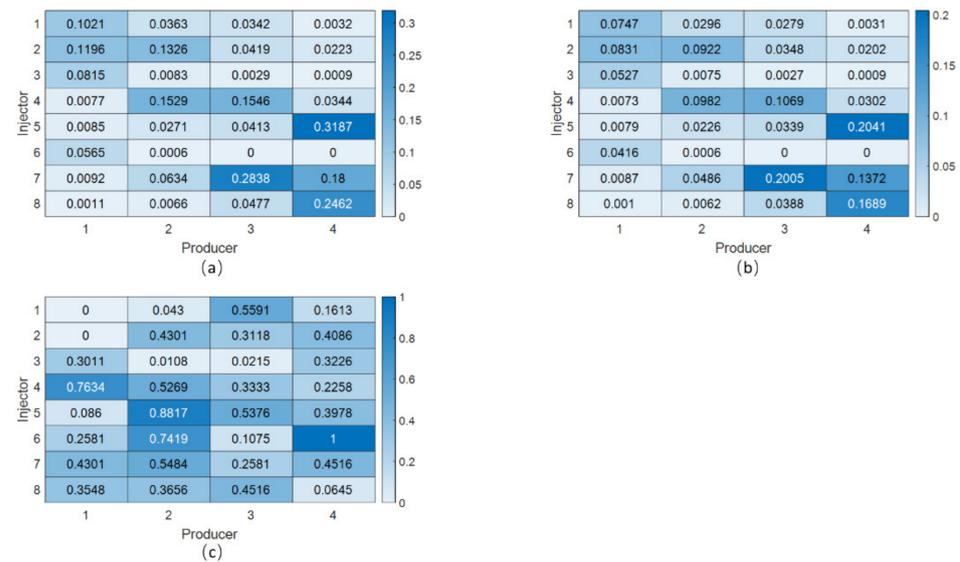


**Figure 15.** The training error (MSE) curves of KINN-tansig, KINN-Gaussian, and SLFNN for the Egg reservoir case. The black line with squares is the error curve of KINN-tansig; the blue line with triangles represents the error curve of KINN-Gaussian, and the green line with stars represents the error curve of SLFNN.

**Table 6.** The time consumption, training error (MSE), and testing error (MSE) of KINN-tansig, KINN-Gaussian, and SLFNN in the Egg reservoir case.

|                                            | KINN-Tansig | KINN-Gaussian | SLFNN    |
|--------------------------------------------|-------------|---------------|----------|
| Computation time (training and testing)    | 0.1282 s    | 0.8539 s      | 0.3361 s |
| Error of history matching (training error) | 0.0022      | 0.0035        | 0.0097   |
| Error of prediction (testing error)        | 0.0171      | 0.0263        | 0.0426   |

Figure 16 reveals the inter-well connectivity characterization results obtained by three models. For P1, the production comes from the contribution of I1, I2, and I3. It is important to note that despite the fault between I3 and P1, the injected water of I3 still could reach P1 by getting round of the fault, as shown in Figure 13b. As shown in Figure 16a,b, I1-P1, I2-P1, and I3-P1 can be reflected correctly by the two proposed models, while I6-P1 is wrongly considered as a relatively high connecting well pair by both models. Moreover, the main contributors of P2 are I2 and I4; the main contributors of P3 are I4 and I7; and the flow of P4 comes from I5, I7, and I8. As shown in Figure 16a,b, all these strong connecting well pairs can be accurately revealed by the deeper color blocks of the heatmaps by the two models. However, SLFNN still struggles to characterize the relative connecting strength. For instance, the top four connectivity values obtained by SLFNN are assigned to I6-P4, I5-P2, I4-P1 and I6-P2, which are actually weak connecting well pairs.



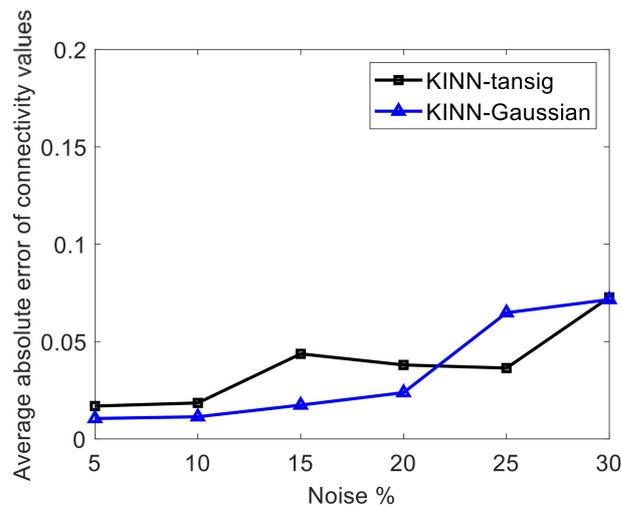
**Figure 16.** The heatmaps of the inter-well connectivity analysis by KINN-tansig and KINN-Gaussian for the Egg reservoir case: (a) KINN-tansig; (b) KINN-Gaussian; (c) SLFNN.

4.4. Sensitivity to Noise

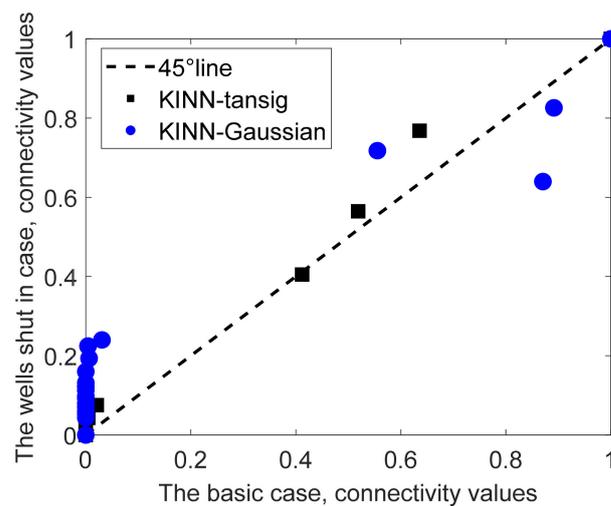
The measurement noise and wells shut-in are unavoidable in real oilfield production, which are common challenges for all reservoir characterization methods. To evaluate the performance of KINN-tansig and KINN-Gaussian on these noisy data, we design the measurement noise case and wells shut-in cases for the braided river reservoir model. In the measurement noise case, all the injection data (5 × 1447 samples) and production data (4 × 1447 samples) are added with Gaussian noises, whose mean value is 1 and standard deviations range from 5%, 10%, 15%, 20%, 25% and 30%, respectively. In the wells shut-in case, 5 injectors and 4 producers are shut in from the time step 401 to 800.

Figure 17 shows the average absolute error of the connectivity values of KINN-tansig and KINN-Gaussian in the measurement noise case. As expected, the error of connectivity values grows slightly with the increase of measurement noise, where the two proposed models show great robustness. As can be seen in Figure 17, the average absolute errors of the connectivity values by the two models are less than 0.08, even though the noise reaches

30%. Figure 18 demonstrates the inter-well connecting relationship by KINN-tansig and KINN-Gaussian in the wells shut-in case against the basic case. KINN-Gaussian shows poor performance, since a part of connectivity values obtained by KINN-Gaussian are higher than 0 in the wells shut-in case, while they are very close to 0 in the basic case. This phenomenon means that the inter-well connectivity values of these weak connecting well pairs obtained by KINN-Gaussian are affected by wells shut-in. In detail, these weak connecting well pairs are likely to get bigger connectivity values in the wells shut-in case than those in the basic case. Nevertheless, KINN-tansig still demonstrates strong robustness in this case. As shown in Figure 18, the connectivity values obtained by KINN-tansig are close to the 45° line, which means that KINN-tansig can generate similar inter-well connectivity characterization results in wells shut-in case as in the basic case.



**Figure 17.** The average absolute error of connectivity values by KINN-tansig and KINN-Gaussian in the noise measurement case. The black line is the error curve of KINN-tansig and the blue line is the error curve of KINN-Gaussian.



**Figure 18.** A cross plot of the connectivity values using KINN-tansig and KINN-Gaussian in the wells shut-in case against the basic case results. The black squares and blue circles represent the connectivity values of KINN-tansig and KINN-Gaussian, respectively, and the dashed is 45° line.

### 5. Discussion and Conclusions

The application of ANN in oil industry is limited by its unexplainability and poor generalizability. In this paper, we concentrate on associating the physical knowledge with neural networks to solve the reservoir characterization and production forecast problems.

Integrating the material balance equation with the machine learning techniques, the physical knowledge interaction neural networks have been proposed, combining both the merits of interpretability and robustness. Furthermore, the proposed gate functions have avoided the negative connectivity values without physical sense, and the computation efficiency has been fully improved by unconstrained optimization algorithm. In the end, the effectiveness of our models has been proved through several simulation experiments. Moreover, the performance of the proposed models on noisy data has been demonstrated. KINN illustrates a novel configuration to realize the cooperation and interaction between neural networks and physical knowledge. In the future, we would like to extend KINN to other areas, like production optimization, and try other machine learning optimization algorithms, like the fractional stochastic gradient descent method [47].

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

**Funding:** This research was funded by the National Natural Science Foundation of China under Grant 51722406, 52074340, and 51874335, the Shandong Provincial Natural Science Foundation under Grant JQ201808, the Fundamental Research Funds for the Central Universities under Grant 18CX02097A, the Major Scientific and Technological Projects of CNPC under Grant ZD2019-183-008, the Science and Technology Support Plan for Youth Innovation of University in Shandong Province under Grant 2019KJH002, the National Science and Technology Major Project of China under Grant 2016ZX05025001-006, 111 Project under Grant B08028, the National Key Research and Development Program of China under Grant 2018AAA0100100, the Fundamental Research Funds for the Central Universities under Grant 20CX05002A and Grant 20CX05012A, and the Source Innovation Scientific and Incubation Project of Qingdao, China under Grant 2020-88.

**Data Availability Statement:** The data used in this paper are not publicly available due to limitations of consent for the original study but could be obtained from Jiang upon the reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Nomenclature

The nomenclature used in this paper is as follows:

| Nomenclature   | Explanations                                                                      |
|----------------|-----------------------------------------------------------------------------------|
| $C_t$          | total compressibility, $\text{bar}^{-1}$                                          |
| $i_k$          | water injection rate, $\text{m}^3/\text{Day}$                                     |
| $J$            | productivity index, $\text{m}^3/\text{Day}/\text{bar}$                            |
| $M$            | number of injectors                                                               |
| $N$            | number of producers                                                               |
| $n$            | time-like variable                                                                |
| $\bar{p}$      | average reservoir pressure, bar                                                   |
| $p_{wf}$       | bottom hole pressure, bar                                                         |
| $\hat{q}$      | estimated production rate, $\text{m}^3/\text{Day}$                                |
| $q_j$          | liquid production rate, $\text{m}^3/\text{Day}$                                   |
| $t$            | time step, Day                                                                    |
| $V_p$          | drainage pore volume, $\text{m}^3/\text{Day}$                                     |
| $\lambda_{kj}$ | inter-well connectivity value                                                     |
| $\gamma_{kj}$  | independent variable of inter-well connectivity of intelligent connectivity model |
| $\rho$         | Pearson correlation coefficient                                                   |
| $\tau_i$       | time constant of capacitance resistance model, Day                                |
| $\Gamma_j$     | comprehensive injection rate, $\text{m}^3/\text{day}$                             |
| $k$            | injector index                                                                    |
| $j$            | producer index                                                                    |

## References

1. Hashan, M.; Jahan, L.N.; Zaman, T.U.; Imtiaz, S.; Hossain, M.E. Modelling of fluid flow through porous media using memory approach: A review. *Math. Comput. Simul.* **2020**, *177*, 643–673. [[CrossRef](#)]
2. Mozolevski, I.; Murad, M.A.; Schuh, L.A. High order discontinuous Galerkin method for reduced flow models in fractured porous media. *Math. Comput. Simul.* **2021**, *190*, 1317–1341. [[CrossRef](#)]
3. Ma, X.; Zhang, K.; Zhang, L.; Yao, C.; Yao, J.; Wang, H.; Jian, W.; Yan, Y. Data-Driven Niching Differential Evolution with Adaptive Parameters Control for History Matching and Uncertainty Quantification. *SPE J.* **2021**, *26*, 993–1010. [[CrossRef](#)]
4. Xu, X.; Wang, C.; Zhou, P. GVRP considered oil-gas recovery in refined oil distribution: From an environmental perspective. *Int. J. Prod. Econ.* **2021**, *235*, 108078. [[CrossRef](#)]
5. Xu, X.; Lin, Z.; Li, X.; Shang, C.; Shen, Q. Multi-objective robust optimisation model for MDVRPLS in refined oil distribution. *Int. J. Prod. Res.* **2021**, *5*, 1–21. [[CrossRef](#)]
6. Yin, F.; Xue, X.; Zhang, C.; Zhang, K.; Han, J.; Liu, B.; Wang, J.; Yao, J. Multifidelity Genetic Transfer: An Efficient Framework for Production Optimization. *SPE J.* **2021**, *26*, 1614–1635. [[CrossRef](#)]
7. Zhang, K.; Zhang, J.; Ma, X.; Yao, C.; Zhang, L.; Yang, Y.; Wang, J.; Yao, J.; Zhao, H. History Matching of Naturally Fractured Reservoirs Using a Deep Sparse Autoencoder. *SPE J.* **2021**, *26*, 1700–1721. [[CrossRef](#)]
8. Xu, X.; Lin, Z.; Zhu, J. DVRP with limited supply and variable neighborhood region in refined oil distribution. *Ann. Oper. Res.* **2022**, *309*, 663–687. [[CrossRef](#)]
9. Heffer, K.J.; Fox, R.J.; McGill, C.A.; Koutsabeloulis, N.C. Novel Techniques Show Links between Reservoir Flow Directionality, Earth Stress, Fault Structure and Geomechanical Changes in Mature Waterfloods. *SPE J.* **1997**, *2*, 91–98. [[CrossRef](#)]
10. Tian, C.; Horne, R.N. Inferring Interwell Connectivity Using Production Data. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dubai, United Arab Emirates, 26–28 September 2016. [[CrossRef](#)]
11. Unal, E.; Siddiqui, F.; Rezaei, A.; Eltaleb, I.; Kabir, S.; Soliman, M.Y.; Dindoruk, B. Use of Wavelet Transform and Signal Processing Techniques for Inferring Interwell Connectivity in Waterflooding Operations. In Proceedings of the SPE Annual Technical Conference and Exhibition, Calgary, AB, Canada, 30 September–2 October 2019. [[CrossRef](#)]
12. Wang, Y.; Kabir, C.S.; Reza, Z. Inferring Well Connectivity in Waterfloods Using Novel Signal Processing Techniques. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dallas, TX, USA, 24–26 September 2018. [[CrossRef](#)]
13. Panda, M.N.; Chopra, A.K. An Integrated Approach to Estimate Well Interactions. In Proceedings of the SPE India Oil and Gas Conference and Exhibition, Society of Petroleum Engineers, New Delhi, India, 17–19 February 1998; p. SPE-39563-MS.
14. Artun, E. Characterizing interwell connectivity in waterflooded reservoirs using data-driven and reduced-physics models: A comparative study. *Neural Comput. Appl.* **2017**, *28*, 1729–1743. [[CrossRef](#)]
15. Jensen, J. Comment on “Characterizing interwell connectivity in waterflooded reservoirs using data-driven and reduced-physics models: A comparative study” by E. Artun. *Neural Comput. Appl.* **2016**, *28*, 1745–1746. [[CrossRef](#)]
16. Yousef, A.A.; Gentil, P.H.; Jensen, J.L.; Lake, L.W. A Capacitance Model To Infer Interwell Connectivity From Production and Injection Rate Fluctuations. *SPE Reserv. Eval. Eng.* **2006**, *9*, 630–646. [[CrossRef](#)]
17. Albertoni, A.; Lake, L.W. Inferring interwell connectivity only from well-rate fluctuations in waterfloods. *SPE Reserv. Eval. Eng.* **2003**, *6*, 6–16. [[CrossRef](#)]
18. Lake, L.W.; Liang, X.; Edgar, T.F.; Al-Yousef, A.; Sayarpour, M.; Weber, D. Optimization Of Oil Production Based On A Capacitance Model Of Production And Injection Rates. In Proceedings of the Hydrocarbon Economics and Evaluation Symposium, Dallas, TX, USA, 1–3 April 2007.
19. Sayarpour, M.; Zuluaga, E.; Kabir, C.S.; Lake, L.W. The use of capacitance–resistance models for rapid estimation of waterflood performance and optimization. *J. Pet. Sci. Eng.* **2009**, *69*, 227–238. [[CrossRef](#)]
20. Sayarpour, M. *Development and Application of Capacitance-Resistive Models to Water/CO<sub>2</sub> Floods*; University of Texas at Austin: Austin, TX, USA, 2008.
21. Mamghaderi, A.; Bastami, A.; Pourafshary, P. Optimization of Waterflooding Performance in a Layered Reservoir Using a Combination of Capacitance-Resistive Model and Genetic Algorithm Method. *J. Energy Resour. Technol.* **2012**, *135*, 013102–013110. [[CrossRef](#)]
22. Zhao, H.; Kang, Z.; Zhang, X.; Sun, H.; Cao, L.; Reynolds, A.C. INSIM: A Data-Driven Model for History Matching and Prediction for Waterflooding Monitoring and Management with a Field Application. In Proceedings of the SPE Reservoir Simulation Symposium, Houston, TX, USA, 23–25 February 2015.
23. Guo, Z.; Reynolds, A.C. INSIM-FT in three-dimensions with gravity. *J. Comput. Phys.* **2019**, *380*, 143–169. [[CrossRef](#)]
24. Guo, Z.; Reynolds, A.C. INSIM-FT-3D: A Three-Dimensional Data-Driven Model for History Matching and Waterflooding Optimization. In Proceedings of the SPE Reservoir Simulation Conference, Society of Petroleum Engineers, Galveston, TX, USA, 10–11 April 2019; p. SPE-193841-MS. [[CrossRef](#)]
25. Zhao, H.; Xu, L.; Guo, Z.; Zhang, Q.; Liu, W.; Kang, X. Flow-Path Tracking Strategy in a Data-Driven Interwell Numerical Simulation Model for Waterflooding History Matching and Performance Prediction with Infill Wells. *SPE J.* **2020**, *25*, 1007–1025. [[CrossRef](#)]
26. Kansao, R.; Yrigoyen, A.; Haris, Z.; Saputelli, L. Waterflood Performance Diagnosis and Optimization Using Data-Driven Predictive Analytical Techniques from Capacitance Resistance Models CRM. In Proceedings of the SPE Europec Featured at 79th EAGE Conference and Exhibition, Paris, France, 12–15 June 2017. [[CrossRef](#)]

27. Olenchikov, D.; Posvyanskii, D. Application of CRM-Like Models for Express Forecasting and Optimizing Field Development. In Proceedings of the SPE Russian Petroleum Technology Conference, Moscow, Russia, 22–24 October 2019.
28. Nguyen, A.P.; Lasdon, L.S.; Lake, L.W.; Edgar, T.F. Capacitance Resistive Model Application to Optimize Waterflood in a West Texas Field. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 30 October–2 November 2011.
29. Guo, Z.; Reynolds, A.; Zhao, H. Waterflooding optimization with the INSIM-FT data-driven model. *Comput. Geosci.* **2018**, *22*, 745–761. [[CrossRef](#)]
30. Chen, B.; Pawar, R.J. Characterization of CO<sub>2</sub> storage and enhanced oil recovery in residual oil zones. *Energy* **2019**, *183*, 291–304. [[CrossRef](#)]
31. Alimohammadi, H.; Rahmanifard, H.; Chen, N. Multivariate Time Series Modelling Approach for Production Forecasting in Unconventional Resources. In Proceedings of the SPE Annual Technical Conference and Exhibition, Virtual. 26–29 October 2020. [[CrossRef](#)]
32. Chen, B.; Harp, D.R.; Lin, Y.; Keating, E.H.; Pawar, R.J. Geologic CO<sub>2</sub> sequestration monitoring design: A machine learning and uncertainty quantification based approach. *Appl. Energy* **2018**, *225*, 332–345. [[CrossRef](#)]
33. Li, Y.; Wang, G.; McLellan, B.; Chen, S.-Y.; Zhang, Q. Study of the impacts of upstream natural gas market reform in China on infrastructure deployment and social welfare using an SVM-based rolling horizon stochastic game analysis. *Pet. Sci.* **2018**, *15*, 898–911. [[CrossRef](#)]
34. Boret, S.E.B.; Marin, O.R. Development of Surrogate models for CSI probabilistic production forecast of a heavy oil field. *Math. Comput. Simul.* **2019**, *164*, 63–77. [[CrossRef](#)]
35. Fumagalli, A.; Zonca, S.; Formaggia, L. Advances in computation of local problems for a flow-based upscaling in fractured reservoirs. *Math. Comput. Simul.* **2017**, *137*, 299–324. [[CrossRef](#)]
36. Xu, X.; Hao, J.; Zheng, Y. Multi-objective Artificial Bee Colony Algorithm for Multi-stage Resource Leveling Problem in Sharing Logistics Network. *Comput. Ind. Eng.* **2020**, *142*, 106338. [[CrossRef](#)]
37. Huang, D.S. Radial Basis Probabilistic Neural Networks: Model and Application. *Int. J. Pattern Recognit. Artif. Intell.* **1999**, *13*, 1083–1101. [[CrossRef](#)]
38. Huang, D.; Du, J. A Constructive Hybrid Structure Optimization Methodology for Radial Basis Probabilistic Neural Networks. *IEEE Trans. Neural Netw.* **2008**, *19*, 2099–2115. [[CrossRef](#)]
39. Karpatne, A.; Watkins, W.; Read, J.; Kumar, V. Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. *arXiv* **2017**, arXiv:1710.11431.
40. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]
41. Tartakovsky, A.M.; Marrero, C.O.; Perdikaris, P.; Tartakovsky, G.D.; Barajas-Solano, D. Physics-Informed Deep Neural Networks for Learning Parameters and Constitutive Relationships in Subsurface Flow Problems. *Water Resour. Res.* **2020**, *56*, e2019WR026731. [[CrossRef](#)]
42. Dehghani, H.; Zilian, A. A hybrid MGA-MSGD ANN training approach for approximate solution of linear elliptic PDEs. *Math. Comput. Simul.* **2021**, *190*, 398–417. [[CrossRef](#)]
43. Wang, N.; Zhang, D.; Chang, H.; Li, H. Deep learning of subsurface flow via theory-guided neural network. *J. Hydrol.* **2020**, *584*, 124700. [[CrossRef](#)]
44. Csiszár, O.; Csiszár, G.; Dombi, J. Interpretable neural networks based on continuous-valued logic and multicriteria decision operators. *Knowl.-Based Syst.* **2020**, *199*, 105972. [[CrossRef](#)]
45. Huang, C.; Xu, H.; Xu, Y.; Dai, P.; Xia, L.; Lu, M.; Bo, L.; Xing, H.; Lai, X.; Ye, Y. Knowledge-aware Coupled Graph Neural Network for Social Recommendation. In Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI), Virtual. 2–9 February 2021. [[CrossRef](#)]
46. Zandvliet, M.J.; Bosgra, O.H.; Jansen, J.-D.; Van den Hof, P.; Kraaijevanger, J.B.F.M. Bang-bang control and singular arcs in reservoir flooding. *J. Pet. Sci. Eng.* **2007**, *58*, 186–200. [[CrossRef](#)]
47. Khan, Z.; Chaudhary, N.I.; Zubair, S. Fractional stochastic gradient descent for recommender systems. *Electron. Mark.* **2019**, *29*, 275–285. [[CrossRef](#)]