applied sciences

MDPI

Article

Prediction of the Productivity Ratio of Perforated Wells Using Least Squares Support Vector Machine with Particle Swarm Optimization

Haijing Wang, Chao Zhang, Bo Zhou *, Shifeng Xue * and Feifei Wang

College of Pipeline and Civil Engineering, China University of Petroleum (East China), Qingdao 266580, China; wanghj@upc.edu.cn (H.W.); s22060013@s.upc.edu.cn (C.Z.); wangfei_upc09@163.com (F.W.)
* Correspondence: zhoubo@upc.edu.cn (B.Z.); sfeng@upc.edu.cn (S.X.)

Abstract: The productivity ratio is a vital metric for assessing the efficiency of perforated completions. Accurate and rapid prediction of this ratio is essential for optimizing the perforation design. In this study, we propose a novel approach that combines three-dimensional finite element numerical simulation and machine learning techniques to predict the productivity ratio of perforated wells. Initially, we obtain the productivity ratio of perforated wells under various perforation parameters using three-dimensional finite element numerical simulation. This generates a sample set for machine learning. Subsequently, we employ the least squares support vector machine (LSSVM) algorithm to establish a prediction model for the productivity ratio of perforated wells. To optimize the parameters of the LSSVM algorithm, we utilize the particle swarm optimization (PSO) algorithm. We compare our proposed PSO-LSSVM model with that established based on other parameter optimization methods and machine learning algorithms, such as Grid search-LSSVM, PSO-ANN, and PSO-RF. Our results demonstrate that the PSO-LSSVM model exhibits rapid convergence, high prediction accuracy, and strong generalization ability in predicting the productivity ratio of perforated wells. This research provides a valuable reference and guidance for optimizing perforation design. Additionally, it offers new insights into predicting the productivity of complex completions.

Keywords: finite element numerical simulation; least squares support vector machine; particle swarm optimization; productivity ratio; perforated wells

check for updates

Citation: Wang, H.; Zhang, C.; Zhou, B.; Xue, S.; Wang, F. Prediction of the Productivity Ratio of Perforated Wells Using Least Squares Support Vector Machine with Particle Swarm Optimization. *Appl. Sci.* **2023**, *13*, 12978. https://doi.org/10.3390/app132412978

Academic Editor: Marek Krawczuk

Received: 4 November 2023
Revised: 1 December 2023
Accepted: 1 December 2023
Published: 5 December 2023

Copyright: © 2023 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

Perforation completion is a prominent technique utilized in the oil and gas industry for completing wells in oil and gas fields. It has gained significant popularity and widespread application in the global petroleum sector. The productivity ratio is a vital metric used to assess the efficiency of perforated completions and serves as a fundamental basis for optimizing the parameters associated with the perforation process. Extensive research has been conducted to investigate the productivity ratio or productivity of wells that have undergone perforation completion [1–5]. Currently, the commonly employed methods for predicting the productivity ratio or productivity of such wells include physical experimental methods, semi-analytical methods, and numerical simulation methods. Physical experimental methods involve artificial rock sample simulation experiments [6,7] and electrical analog experiments [8]. However, these methods have drawbacks such as implementation difficulties, long cycles, and high costs. Semi-analytical methods [9,10] derive formulas based on various assumptions, which often include empirical coefficients related to perforation parameters. These coefficients are challenging to obtain onsite or accurately predict, resulting in low prediction accuracy. On the other hand, numerical simulation methods [11–15] can replicate the physical processes of reservoir exploitation but require extensive data on oilfield geology and development, which are often difficult to obtain or have significant uncertainties. This limitation severely impacts the accuracy of a

numerical simulation. Additionally, numerical simulation methods demand substantial computational resources, have long processing times, low efficiency, and necessitate prediction personnel with a solid foundation in numerical simulation, posing inconveniences for engineering applications. Some researchers [16,17] have performed regression analysis on productivity ratio data obtained from numerical simulation and derived simple linear or nonlinear fitting formulas. However, due to the numerous influencing factors and complex relationships involved in the productivity ratio, these fitting formulas obtained through simple regression often exhibit high errors and limited applicability. Consequently, there is an urgent need to explore a simpler, more efficient, and data-driven method for predicting the productivity ratio of perforated wells.

In recent years, machine learning techniques have rapidly developed in the field of petroleum engineering. Their applications involve reservoir characterization and evaluation [18,19], lithology classification and geological modeling [20,21], reservoir fluid property prediction [22,23], production optimization and intelligent decision making [24,25], and fault prediction and maintenance [26,27], providing new approaches for the productivity ratio prediction of perforated wells. Commonly used machine learning regression algorithms include artificial neural networks (ANN) [28–30], support vector machines (SVM) [31–33], decision trees (DT) [34–36], and random forests (RF) [37–39]. However, these methods have certain issues. ANN often require long training times, especially when dealing with large-scale datasets. Moreover, they and algorithms like DT tend to have high model complexity, which makes them susceptible to overfitting the training set and, consequently, leads to decreased performance on the test set. Additionally, these algorithms have multiple hyperparameters that need to be adjusted, making model optimization a complex task. Although support vector machines can overcome some of these drawbacks, they also have limitations in practical applications, such as the presence of noise in training data and the inability to handle nonlinear problems. In contrast, least squares support vector machine (LSSVM) offers several advantages. It has a relatively fast training speed, simpler hyperparameter settings, strong regularization ability, better ability to handle overfitting problems, and improved performance with noisy data and nonlinear problems.

The objective of this study is to develop a productivity ratio prediction model for perforated wells by integrating the particle swarm optimization (PSO) with the least squares support vector machine (LSSVM). Specifically, the LSSVM technique is employed to establish the productivity ratio prediction model for perforation wells, while the PSO algorithm is utilized to optimize the parameters within the LSSVM algorithm. In comparison to other parameter optimization methods and machine learning algorithms, the combination of the PSO algorithm and the LSSVM technique proves to be a viable approach for predicting the productivity ratio of perforated wells.

2. Three-Dimensional Finite Element Analysis for Evaluating the Productivity Ratio of Perforated Wells

The productivity ratio (PR) is defined as the ratio of the steady-state productivity of a perforated well to the steady-state productivity of an ideal open-hole well under the same conditions. The mathematical expression is as follows:

$$\mathrm{PR} = \frac{Q}{Q_0}, \tag{1}$$

where PR is the productivity ratio of the perforated well, Q is the productivity of the perforated well, m^3/d, and Q_0 is the productivity of the ideal open-hole well, m^3/d, which can be calculated using the following theoretical formula:

$$Q_0 = \frac{2\pi Kh(p_e - p_w)}{\mu ln r_w/r_e}, \tag{2}$$

where K is the formation permeability, h is the formation thickness, p_w and p_e are the bottom hole pressure and external boundary pressure, respectively, μ is the fluid viscosity, and r_w and r_e are the radius of the wellbore and the circular formation, respectively.

Due to the asymmetry of the perforation hole arrangement and the complexity of the flow paths, the productivity of perforated wells cannot be accurately determined using analytical methods. Therefore, numerical methods, specifically the finite element method, are employed in this study to calculate the productivity of perforated wells under varying perforation parameters.

2.1. Finite Element Model

When calculating productivity, it is assumed that the fluid flow within the reservoir is isothermal and in a steady state, adhering to Darcy's law. The flow equation can be expressed as follows:

$$\frac{\partial}{\partial x}\left(K_x\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial p}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial p}{\partial z}\right)=0, \tag{3}$$

where p is the fluid pressure, Pa; and K_x, K_y, and K_z are the permeabilities in the x, y, and z directions, respectively, m^2.

Due to the heterogeneity of the formation, the size and shape of the perforation holes created by the same type of perforating gun may not be identical. Additionally, the fragments generated during perforation can obstruct the holes, resulting in an unevenly compacted zone around the holes. These uncertainties pose challenges in modeling perforated wells. To simplify the modeling process and facilitate subsequent analysis, the following simplifications and assumptions are made regarding the perforation holes: (1) Perforation holes created by the same type of perforating gun are assumed to be identical. (2) The geometric shape of the perforation holes is considered to be a regular cylinder. (3) All the holes are assumed to be unobstructed, and the thickness and compaction degree of the compacted zone around the holes are assumed to be uniform. (4) The shape of the drilling mud-contaminated zone is assumed to be a cylindrical annulus, uniformly surrounding the wellbore, and the degree of drilling mud contamination is also assumed to be uniform. Based on these assumptions, a 3D geometric model of a perforated well, encompassing the wellbore, perforation holes, compacted zone, and contamination zone, is established as depicted in Figure 1.

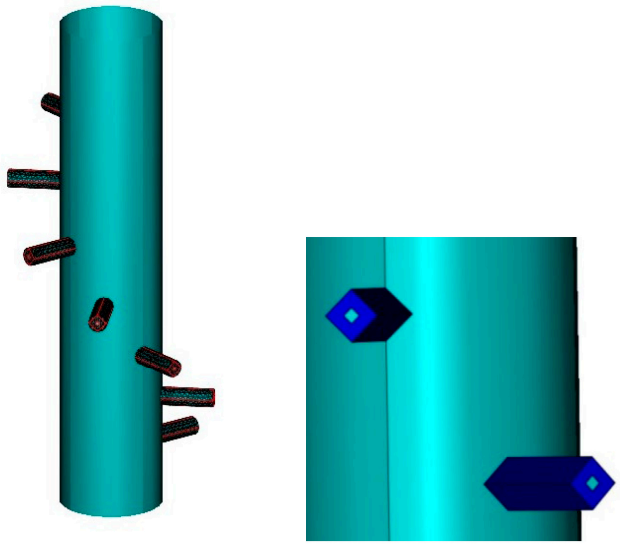


Figure 1. Geometric model of a perforated well.

The mesh partitioning of geometric models in the context of wellbore simulation requires careful consideration of both the computational accuracy and efficiency. In this study, a two-step approach is employed to achieve an optimal mesh refinement. Firstly, an adaptive mesh refinement method is utilized to refine the mesh around the wellbore region, taking into account the significant difference in size between the formation and the wellbore. Subsequently, a mesh size sensitivity analysis is conducted to determine the final mesh partitioning method, as illustrated in Figure 2. The maximum element size is set as follows: 0.005 m for the compacted zone around the perforation hole, 0.02 m for the contaminated zone around the wellbore, 0.05 m for the near-wellbore formation region, and 0.5 m for the peripheral formation region.

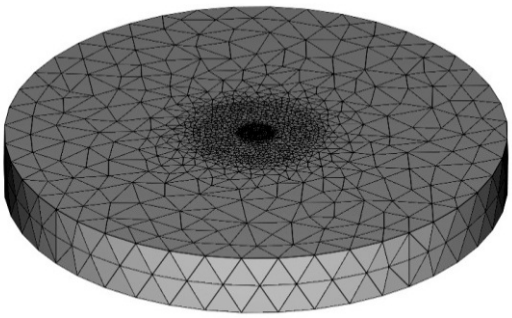


Figure 2. Mesh grids.

To simulate the flow field of the reservoir during production from the perforated well, pressure boundary conditions are applied. Specifically, a pressure difference of 1 MPa is assumed, with 1 MPa and 2 MPa applied as pressure boundary conditions to the inner wall of the wellbore and the outer boundary of the reservoir, respectively. The streamline distribution around the wellbore obtained from the numerical simulation is shown in Figure 3. The results indicate that the reservoir fluid primarily enters the wellbore through the perforation holes. By integrating the flow rate at each perforation boundary, the productivity Q of the well can be determined. Additionally, by substituting the same parameters into Equation (2), the productivity of an ideal open-hole well without pollution and perforations can be calculated. Finally, Equation (1) is employed to calculate the productivity ratio of the perforated well. The accuracy of the aforementioned finite element numerical simulation methods was validated by comparing the productivity ratio data with reference [7] under the given parameters.

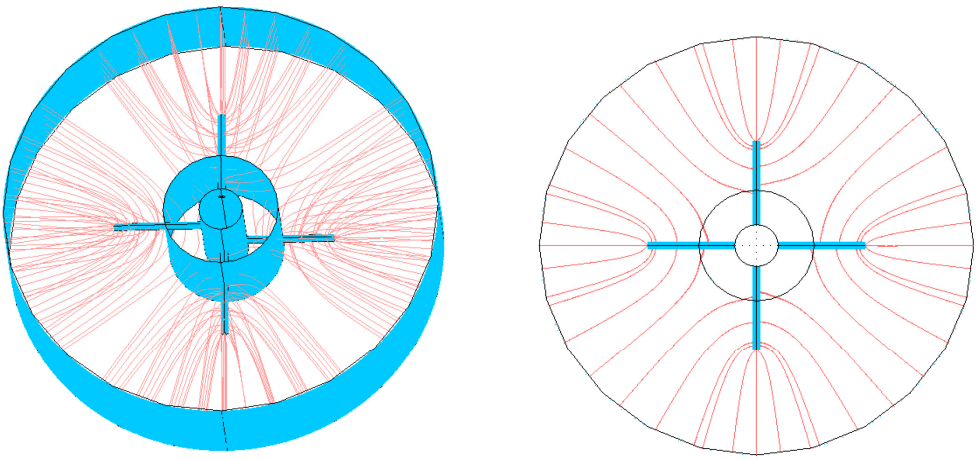


Figure 3. Streamline distribution around the wellbore.

2.2. Influence of the Perforation Parameters

The impact of the perforation parameters on the productivity ratio of oil wells was analyzed by conducting numerical simulations, as described above.

2.2.1. Perforation Depth and Shot Density

As shown in Figure 4, the productivity of perforated wells increases with the perforation depth. When the perforation hole does not penetrate the contaminated zone, the productivity ratio increases linearly with the perforation depth, showing a stable growth rate. When the perforation hole just penetrates the contaminated zone, the productivity ratio increases significantly. However, when the perforation hole extends beyond the contaminated zone by approximately 50 mm, the growth rate of the productivity ratio gradually decreases. Furthermore, the productivity ratio increases with shot density, but the growth rate also decreases gradually. Particularly, when the shot density exceeds 24 holes/m, the influence of increasing shot density on enhancing the productivity ratio becomes less significant.

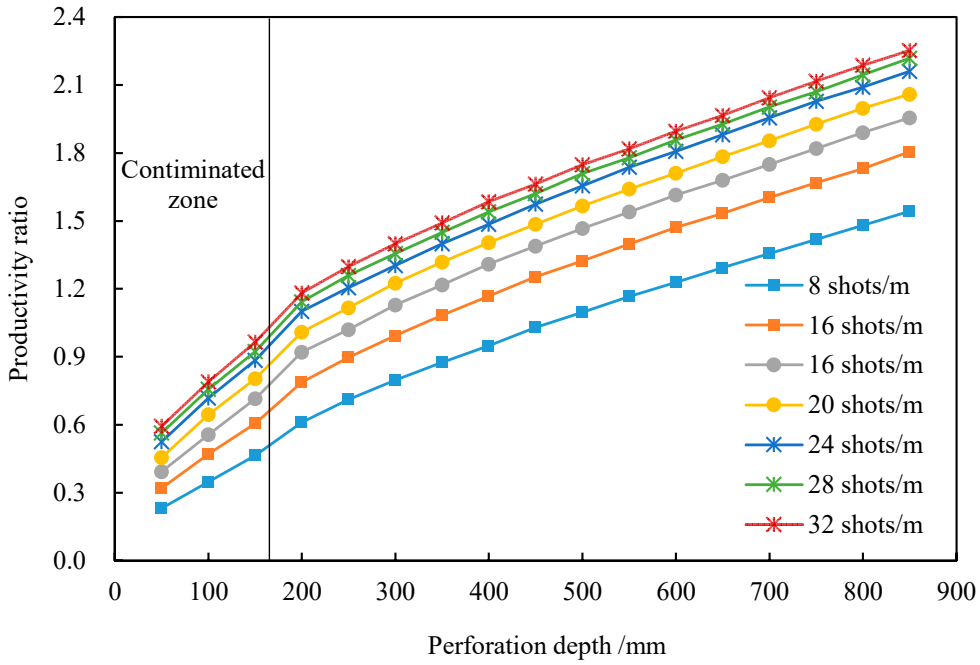


Figure 4. Influence of the perforation depth and shot density on the productivity ratio.

2.2.2. Perforation Phasing

The productivity ratio curves under four commonly used phase angles (0°, 90°, 120°, 180°) in engineering are shown in Figure 5. It can be observed that in isotropic formations, the productivity ratio is highest at 90° and lowest at 0°. The phase angles are ranked in order of decreasing productivity ratio as follows: 90°, 120°, 180°, 0°. This trend can be attributed to the more uniform distribution of perforation holes in the circumferential direction of the wellbore at lower phase angles, such as 90°, when the shot density remains constant.

Furthermore, there is an interaction between phase angles and formation anisotropy. Figure 6 shows the influence of phase angles and formation anisotropy on the productivity ratio when the perforation hole penetrates the drilling contaminated zone. In cases where the anisotropy is not severe ($0.3 < K_v/K_h \leq 1.0$), the phase angle of 90° exhibits the highest performance, while the phase angle of 0° demonstrates the lowest. The phase angles can be ranked in decreasing order of productivity ratio as follows: 90°, 120°, 180°, 0°. Conversely, when the anisotropy is severe ($K_v/K_h \leq 0.2$), the phase angles can be ranked in decreasing order of productivity ratio as follows: 180°, 120°, 90°, 0°. This indicates that higher phase angles result in higher productivity ratios. This phenomenon can be attributed to the

reduction in longitudinal spacing between adjacent perforation holes in the same direction at higher phase angles, thereby minimizing the impact of anisotropy to the greatest extent possible, when the shot density remains constant.

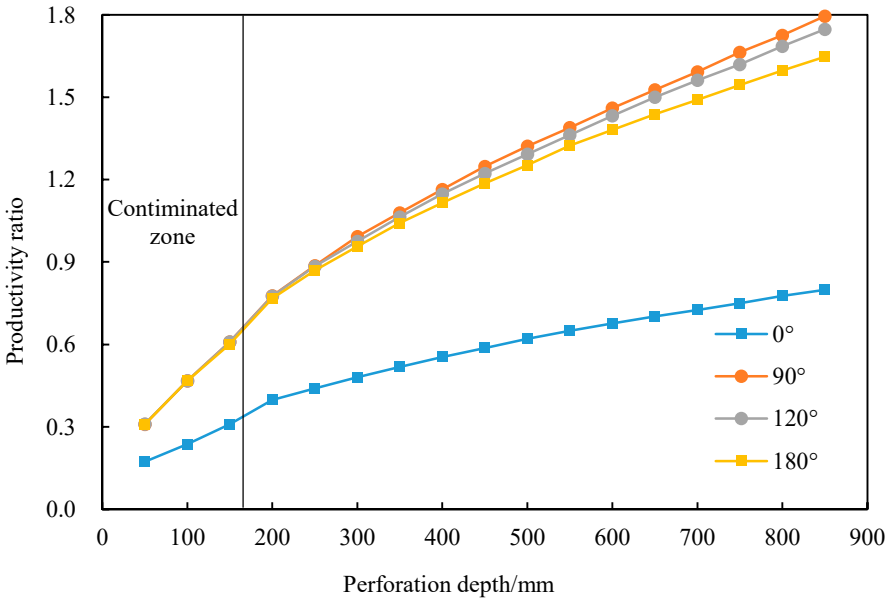


Figure 5. Influence of the phase angles on the productivity ratio in isotropic formations.

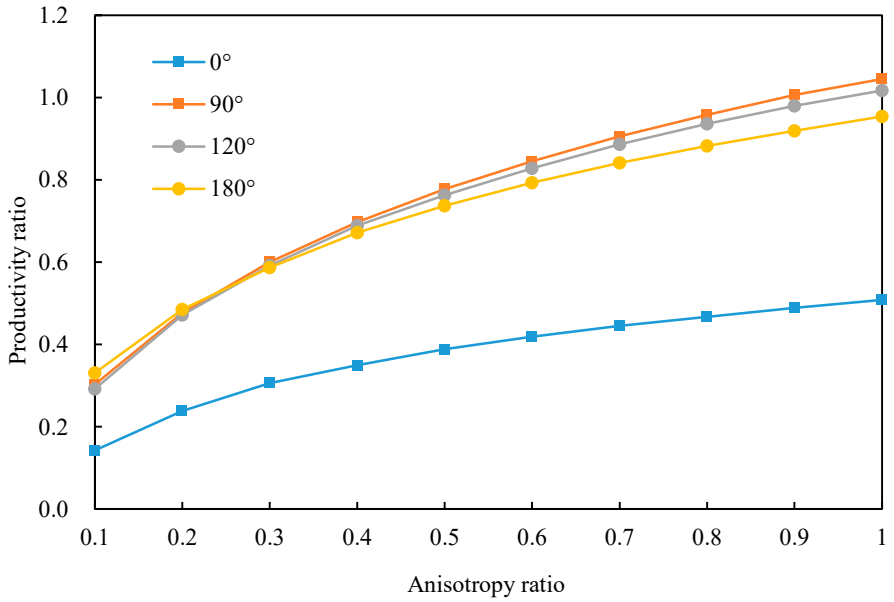


Figure 6. Influence of the phase angles and formation anisotropy on the productivity ratio.

2.2.3. Perforation Damage

During the process of jet-assisted perforating, high-temperature and high-pressure shock waves are generated around the perforation hole, forming a compacted zone with a permeability significantly lower than that of the original formation. The compaction damage process of the formation by shock waves is highly complex. To simplify the calculations and analysis, this study focused solely on simulating the impact of the perforation compaction degree on the productivity ratio, rather than simulating the physical mechanics of the perforation process. In perforated completions, the degree of the perforation compaction damage is typically quantified by the permeability reduction factor, which is defined as the ratio of the permeability in the compacted zone to the original permeability.

The impact of compaction damage on the productivity ratio is illustrated in Figure 7. It can be observed that the compaction damage has a noticeable effect on the productivity ratio. When other parameters remain constant, the productivity ratio decreases as the compaction damage increases (i.e., the permeability reduction factor decreases). The influence of the compaction damage on the productivity ratio becomes even more significant after the perforation hole penetrates the drilling mud-contaminated zone.

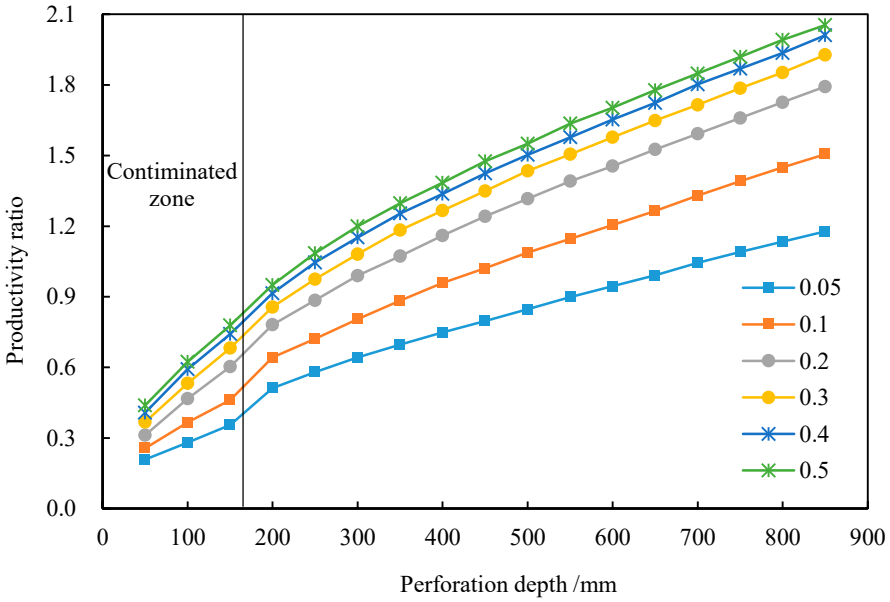


Figure 7. Influence of the perforation compaction damage on the productivity ratio.

3. Principle of Least Squares Support Vector Machine (LSSVM)

The Support Vector Machine (SVM) algorithm is a statistical learning theory developed for situations with limited samples based on the principle of minimizing structural risk. However, the traditional SVM approach requires solving a quadratic programming problem, which is often slow and suffers from the curse of dimensionality, resulting in high computational complexity.

In contrast, the LSSVM [40] differs from the standard SVM in that the loss function of the optimization objective is represented by the Euclidean norm of the error. It converts the inequality constraints in SVM into equality constraints, thereby transforming the training process from solving a quadratic programming problem to solving a system of linear equations. This transformation simplifies the computational complexity and improves the convergence speed.

Consider a training dataset $(\boldsymbol{x}_i, \boldsymbol{y}_i)$, i =1, 2, …, m, where $\boldsymbol{x}_i$ is the i-th input vector, $\boldsymbol{y}_i$ is the corresponding output data, and m represents the number of samples. By mapping the input vectors to a high-dimensional feature space, it is possible to achieve linear separability of the samples in the linear space. Let $\varphi(\boldsymbol{x})$ represent the mapped feature vector of $\boldsymbol{x}$. Consequently, the model corresponding to the hyperplane dividing the feature space can be expressed as:

$$f(\boldsymbol{x}) = \boldsymbol{\omega}^T \varphi(\boldsymbol{x}) + b, \tag{4}$$

where $\boldsymbol{\omega}^T = (\omega_1, \omega_2, \ldots, \omega_n)$ is the normal vector that determines the direction of the hyperplane, and b is the displacement term, a scalar constant that determines the distance between the hyperplane and the origin.

According to the principle of Structural Risk Minimization (SRM) [41], the optimization problem is transformed into:

$$\begin{gathered} \min F(\boldsymbol{\omega}, b, \boldsymbol{e}) = \frac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\omega} + \frac{1}{2}C\sum_{i=1}^{m} e_i^2 \\ \text{s.t. } y_i = \boldsymbol{\omega}^T\varphi(\boldsymbol{x}) + b + e_i, \;\; i = 1, 2, \ldots, m, \end{gathered} \tag{5}$$

where $\boldsymbol{e} = [e_1, e_1, \ldots, e_m]$ is the deviation vector reflecting the regression error between the actual and predicted values, and C is the regularization parameter used to control the penalty for samples exceeding the calculation error, which is manually set.

By introducing the Lagrange function and solving according to the Karush–Kuhn–Tucker (KKT) conditions, the regression function model of the LS-SVM can be obtained:

$$y(\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b, \tag{6}$$

where α_i is the Lagrange multiplier, subject to the constraint $0 \leq \alpha_i \leq C$; $K(\cdot, \cdot)$ is the kernel function, mainly used for inner product operations on sample data.

Therefore, when using LSSVM to establish a regression model, it is crucial to appropriately select the regularization parameter C and the kernel function parameters.

4. Principle of Particle Swarm Optimization (PSO)

The Particle Swarm Optimization (PSO) algorithm is a parameter optimization algorithm based on swarm intelligence. Compared to traditional parameter optimization methods, the PSO algorithm is simple to implement and has high optimization efficiency. It greatly reduces the optimization time and improves the convergence speed through information sharing and transmission. It was first proposed by J. Kennedy and R. C. Eberhart [42] in 1995, inspired by the study of bird flocking behavior. In the PSO algorithm, each solution in the objective space is represented by a bird (particle), and the optimal solution of the problem is analogous to the food source that the bird flock seeks. During the process of finding the optimal solution, each particle exhibits individual behavior as well as collective behavior. Each particle learns from the flying experiences of its companions and incorporates its own flying experiences to locate the optimal solution. Two values are learned by each particle: the individual's historical best value and the group's historical best value. The particle adjusts its velocity and position based on these two values, and the quality of each position is determined by the fitness value. The flowchart of the PSO algorithm is depicted in Figure 8.

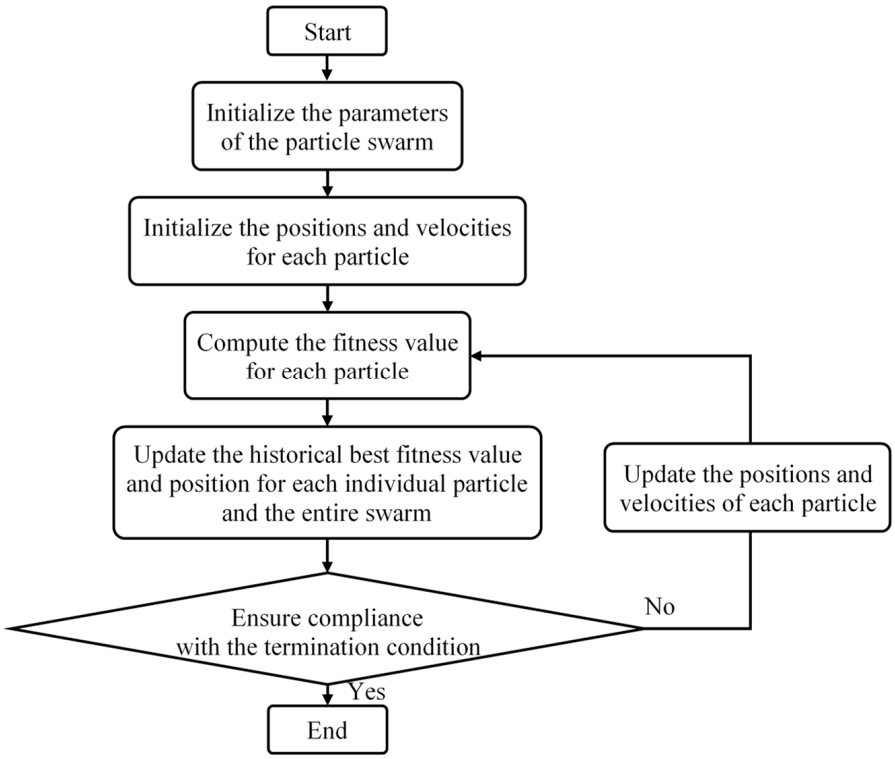


Figure 8. Flowchart of the PSO algorithm.

The detailed parameter optimization steps are as follows:

- **Step 1:** Initialize the parameters of the particle swarm, such as the swarm size, particle dimension, number of iterations, inertia weight, etc.
- **Step 2:** Initialize the position and velocity of each particle.

In a D-dimensional target search space, a particle swarm consists of N particles, where each particle is a D-dimensional vector, and its spatial position represents a solution in the target optimization problem, denoted as

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{iD}), i = 1, 2, \ldots, N. \tag{7}$$

The flying velocity of the i-th particle is also a D-dimensional vector, denoted as

$$v_i = (v_{i1}, v_{i2}, \ldots, v_{iD}), i = 1, 2, \ldots, N. \tag{8}$$

The positions and velocities of particles are randomly generated within a given range.

- **Step 3:** Calculate the fitness value based on the fitness function. The fitness function represents the objective function of the optimization problem, which can be the mean square error, variance, standard deviation, etc.
- **Step 4:** Update the historical best fitness value and position for each individual and the entire swarm.

For each particle, compare its fitness value with its historical best fitness value. If the current fitness value is better, update it as the individual's historical best value. Then, compare its fitness value with the fitness value of the best position experienced by the entire swarm. If the current fitness value is better, update it as the swarm's historical best value.

- **Step 5:** Determine if the termination condition is met (reaching the maximum number of iterations or the minimum difference in fitness value between two consecutive iterations). If the termination condition is met, output the optimal position. Otherwise, update the position and velocity of each particle and repeat steps 3–4.

The velocity update formula is as follows:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_1 \Big(p_{bestij}(t) - x_{ij}(t) \Big) + c_2 r_2 \Big(g_{bestj} - x_{ij}(t) \Big), \tag{9}$$

where $v_{ij}(t)$ is the velocity of particle i in the j-th dimension at time t. The subscript i represents the i-th particle, the subscript j represents the j-th dimension of the particle, which corresponds to the j-th parameter optimized by the algorithm, and t represents the current iteration count. $c1$ and $c2$ are acceleration constants, usually taken in the range of 0 to 2. $r1$ and $r2$ are two independent random numbers that vary between 0 and 1. $p_{bestij}(t)$ is the historical best solution of particle i in the j-th parameter at time t, and g_{bestj} is the historical best solution of all particles in the j-th parameter.

The position update formula is as follows:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \tag{10}$$

where $x_{ij}(t)$ is the position of particle i in the j-th dimension at time t.

5. Prediction Model of Productivity Ratio for Perforated Wells

5.1. Data Preparation

This work utilizes numerical calculation data as sample data for the prediction model. Firstly, five perforation parameters, namely, the perforation depth, shot density, perforation diameter, thickness, and permeability reduction coefficient of the compacted zone, are varied (as shown in Table 1) to generate the original dataset consisting of 4^5 = 1024 datasets. Then, the original data are normalized using linear function normalization technique to enhance the performance and accuracy of the model.

Table 1. Input parameters of the numerical model.

Parameter	Value
Perforation depth (mm)	100, 300, 500, 700
Perforation diameter (mm)	8, 10, 12, 14
Shot density (shots/m)	8, 16, 24, 32
Thickness of the compacted zone (mm)	10, 12, 14, 16
Permeability reduction coefficient of the compacted zone	0.1, 0.2, 0.3, 0.4

5.2. Model Establishment

In this study, LSSVM is employed to establish the prediction model of the productivity ratio for perforated wells. Due to the nonlinearity of the data, the Gaussian kernel function is selected as the kernel function of the LSSVM algorithm, which represents a nonlinear mapping relationship as follows:

$$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right), \tag{11}$$

where σ denotes the bandwidth of the Gaussian kernel. The Gaussian kernel function is widely adopted as a universal kernel function, applicable to samples with any distribution as long as the selected kernel parameter σ is appropriate.

To establish the LSSVM model, reasonable choices for the penalty factor c and the kernel width σ are crucial. Traditional methods often rely on trial-and-error or grid search algorithms. The trial-and-error method requires practical application experience and algorithm analysis ability, while the grid search method is closely related to the optimization step size. If the step size is too large, it may fail to obtain the global optimal solution, and if the step size is too small, it will be time consuming. To establish the optimal model, this study adopts the PSO algorithm to optimize and select the two parameters of the LSSVM algorithm. The dimension of the particle swarm is set to 2, the number of particles in each dimension of the particle swarm is set to 10, and the iteration times are set to 100. The fitness function is set as the Relative Root Mean Square Error (RRMSE) of the model prediction results.

5.3. Results Analysis

Based on the aforementioned model, the PSO is employed to optimize the parameters of the LSSVM in the prediction of the productivity ratio for perforated wells. The PSO error curve is depicted in Figure 9, where the horizontal axis represents the epoch, and the vertical axis represents the test set RRMSE. The figure demonstrates that the PSO algorithm achieves a relatively ideal solution in less than 10 steps, while the global optimal solution is attained in approximately 90 steps, indicating rapid convergence. The total time consumed is approximately 11 h, 44 min, and 56.3163 s. The optimized values for the parameters c and σ are determined as 500.428 and 0.554, respectively. By substituting these optimized parameters into the LSSVM model for predicting the productivity ratio, the test set RRMSE is found to be 0.0004567.

For comparison, the grid search algorithm is also employed to optimize the parameters of the LSSVM. The penalty factor C is selected within the optimization range of 0.001 to 1000, with a step size of 1, while the kernel width σ is selected within the optimization range of 0.001 to 10, with a step size of 0.5. The entire optimization process requires a total of 20,000 steps to complete, taking approximately 24 h, 26 min, and 54.0125 s. The optimized values for c and σ are determined as 500.001 and 0.501, respectively. By utilizing these obtained parameters to reconstruct the productivity ratio prediction model, the test set RRMSE is found to be 0.0005501.

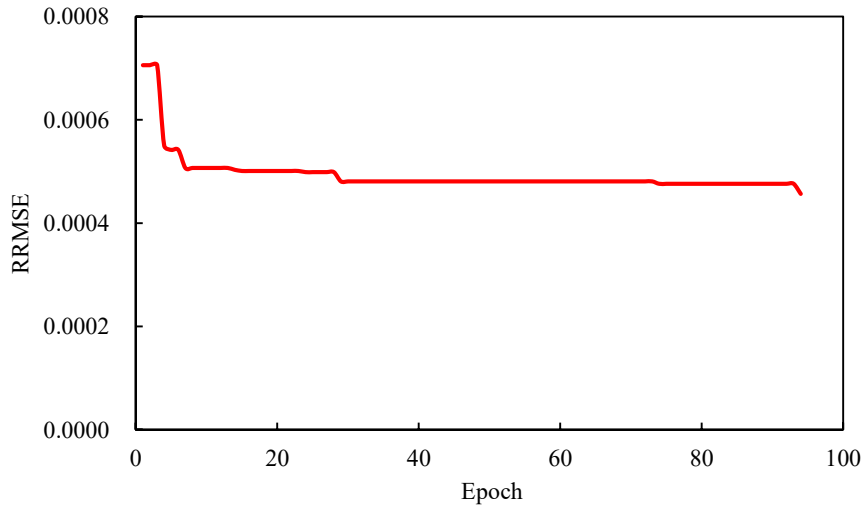


Figure 9. Process of optimization based on the PSO.

Table 2 presents the time, test set RRMSE, and final model parameters for predicting the productivity ratio using the two optimization algorithms. It is evident that the PSO algorithm outperforms the grid search algorithm in terms of computational efficiency and accuracy.

Table 2. Comparison of results between the PSO and grid search.

Algorithm	Time	RRMSE	Parameters	
			C	σ
PSO	11:44:56.3163	0.0004567	500.428	0.554
Grid Search	24:26:54.0125	0.0005501	500.001	0.501

In order to determine the optimal predictive model for the productivity ratio of perforated wells, this study also combines PSO with an Artificial Neural Network (ANN) and Random Forest (RF) to establish corresponding machine learning models. In addition to the RRMSE, the R2 (determination coefficient) and MAPE (mean absolute percentage error) are also used as statistical evaluation indicators for the three machine learning models, as shown in Table 3 and Figure 10. It is evident that among the three models, the machine learning model based on PSO-LSSVM performs the best.

Table 3. Statistical evaluation metrics of the three machine learning models.

Method	RRMSE	Score	R2	Score	MAPE	Score	Total Score
PSO-LSSVM	0.00004567	3	0.9948	3	0.00156	3	9
PSO-ANN	0.01826	1	0.9735	1	0.00836	2	4
PSO-RF	0.002456	2	0.9896	2	0.01049	1	5

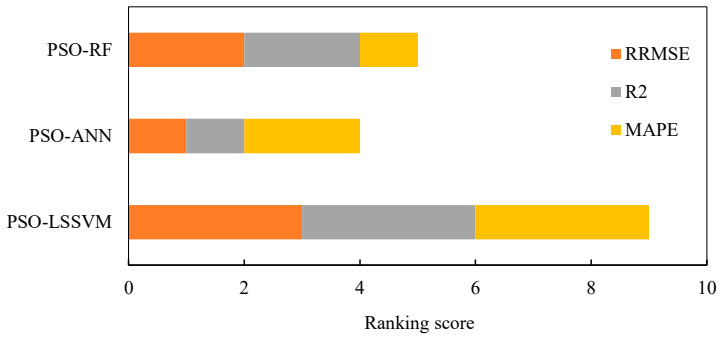


Figure 10. Comparison of the ranking scores of the three machine learning models.

The comparison of the predicted values and actual values for the productivity ratio based on the PSO-LSSVM, is presented in Figure 11. It is evident that the predicted values closely match the actual values, indicating a high predictive capability of the PSO-LSSVM model. Therefore, the combination of PSO and LSSVM to establish a predictive model for the productivity ratio of perforated wells is proven to be effective and feasible.

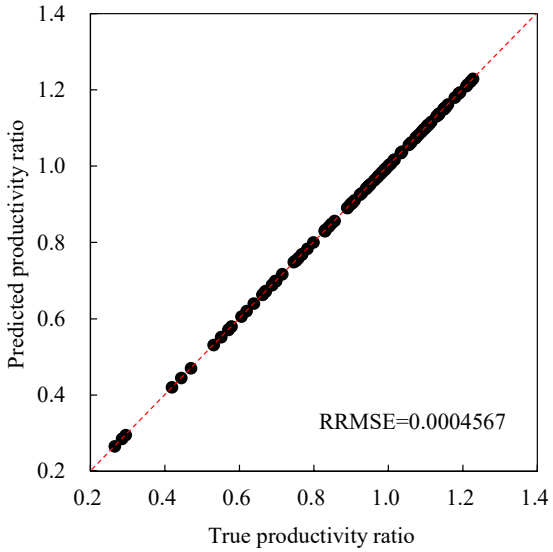


Figure 11. Comparison of the predicted values and actual values based on PSO-LSSVM.

6. Conclusions

This study proposes a predictive method for the productivity ratio of perforated wells based on three-dimensional finite element numerical simulation and machine learning technology. By utilizing the numerical simulation results as the sample data for machine learning, a predictive model for the productivity ratio of perforated well is established based on the LSSVM. The optimal parameters of the LSSVM predictive model are obtained using PSO. Compared to the traditional grid search algorithm, the PSO algorithm significantly improves the modeling efficiency, reducing the time required by more than half. Furthermore, compared to the PSO-ANN and PSO-RF models based on ANN and RF, the predictive accuracy and generalization performance of the productivity ratio model established based on PSO-LSSVM are higher, meeting the requirements of engineering practice. Therefore, the combination of PSO and LSSVM for predicting the productivity ratio of perforated well is proven to be effective and feasible. This research holds important reference value and guiding significance for perforation optimization design and also provides new insights for predicting the productivity of complex completions and geothermal wells.

Author Contributions: Conceptualization, H.W.; methodology, B.Z. and S.X.; software, H.W.; validation, H.W., C.Z. and F.W.; writing—original draft preparation, H.W.; writing—review and editing, C.Z.; visualization, H.W. and C.Z.; supervision, B.Z. and S.X.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the support from the National Natural Science Foundation of China, grant number 51804323.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding authors, Bo Zhou and Shifeng Xue, upon reasonable request. The data are not publicly available due to privacy.

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

References

1. Xing, X.; Qiu, H.; Wen, M.; Hou, Z.-N.; Huang, H.; Bi, G. Numerical Simulation Investigation of Perforation Parameter Optimization Design Based on Orthogonal Experimental Method. In Proceedings of the International Field Exploration and Development Conference, Xi'an, China, 16–18 November 2022; Lin, J., Ed.; Springer Series in Geomechanics and Geoengineering; Springer: Singapore, 2023.
2. Wu, M.; Zhu, J.; Li, L.; Li, P. Calculation of Perforated Vertical and Horizontal Well Productivity in Low-Permeability Reservoirs. *SPE Drill Compl.* **2020**, *35*, 218–236. [CrossRef]
3. Panteha, G.; Mahmoud, J.; Ebrahim, A.; Wilkinson, D.; Dehkordi, F.S.; Hamidi, H. A new and simple model for the prediction of horizontal well productivity in gas condensate reservoirs. *Fuel* **2018**, *223*, 431–450.
4. Grove, B.; Harvey, J.; Martin, A. Perforating System Performance at Downhole Conditions: Recent Advances in Modeling and Prediction. In Proceedings of the SPE International Conference and Exhibition on Formation Damage Control, Lafayette, LA, USA, 24–26 February 2016.
5. Roostaei, M.; Soroush, M.; Mohammadtabar, F.; Mohammadtabar, M.; Hosseini, S.A.; Mahmoudi, M.; Sadrzadeh, M.; Ghalambor, A.; Fattahpour, V. Design for Reliability: Experimental and Numerical Simulation of Cased and Perforated Completions with Standalone Screen. *SPE Drill Compl.* **2021**, *36*, 680–706. [CrossRef]
6. Allison, S.; Löhken, J.; Nelis, L.; Djayapertapa, L.; Byrne, M.; Watson, K.; Clarke, A. Understanding Perforation Geometry Influence on Flow Performance Using CFD. In Proceedings of the SPE European Formation Damage Conference and Exhibition, Budapest, Hungary, 3–5 June 2015.
7. Abobaker, E.; Elsanoose, A.; Khan, F.; Rahman, M.A.; Aborig, A.; Noah, K.; Butt, S. A new assessment of perforation skin factor for vertical perforated wells in near-wellbore region. *J. Petrol. Explor. Prod. Technol.* **2022**, *12*, 117–133. [CrossRef]
8. Yildiz, T.; Ozer, D. Experimental Study on the Productivity of Complex Well Configurations. In Proceedings of the SPE International Conference on Horizontal Well Technology, Calgary, AB, Canada, 1–4 November 1998.
9. Karakas, M.; Tariq, S. Semianalytical Productivity Models for Perforated Completions. *SPE. Prod. Oper.* **1991**, *6*, 73–82. [CrossRef]
10. Yildiz, T. Assessment of Total Skin Factor in Perforated Wells. *SPE Res. Eval. Eng.* **2006**, *9*, 61–76. [CrossRef]
11. Locke, S. An Advanced Method for Predicting the Productivity Ratio of a Perforated Well. *J. Pet. Technol.* **1981**, *33*, 2481–2488. [CrossRef]
12. Rahman, M.; Mustafiz, S.; Koksal, M.; Islam, M. Quantifying the skin factor for estimating the completion efficiency of perforation tunnels in petroleum wells. *J. Petrol. Sci. Eng.* **2007**, *58*, 99–110. [CrossRef]
13. Sun, D.; Li, B.; Gladkikh, M.; Satti, R.; Evans, R. Comparison of Skin Factors for Perforated Completions Calculated with Computational-Fluid-Dynamics Software and the Karakas-Tariq Semianalytical Model. *SPE Drill Compl.* **2013**, *28*, 21–33. [CrossRef]
14. Abobaker, E.; Elsanoose, A.; Khan, F.; Rahman, M.A.; Aborig, A.; Noah, K. Quantifying the partial penetration skin factor for evaluating the completion efficiency of vertical oil wells. *J. Pet. Explor. Prod. Technol.* **2021**, *11*, 1–13. [CrossRef]
15. Zhang, Z.; Guo, J.; Liang, H.; Liu, Y. Numerical simulation of skin factors for perforated wells with crushed zone and drilling-fluid damage in tight gas reservoirs. *J. Nat. Gas. Sci. Eng.* **2021**, *90*, 103907. [CrossRef]
16. Tang, Y.; Ozkan, E.; Kelkar, M.; Yildiz, T. Fast and Accurate Correlations to Compute the Perforation Pseudoskin and Productivity of Horizontal Wells. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September–3 October 2001.
17. Bi, G.; Han, F.; Li, M.; Wu, J.; Cui, Y.; Wang, X. Research on Productivity Prediction Model of Three-Dimensional Directional Wells in Different Reservoirs. *J. Energ. Eng.* **2023**, *149*, 04023020. [CrossRef]
18. Daniel, A.; Tarek, O.; Raoof, G.; Ridha, S. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *J. Petrol. Sci. Eng.* **2021**, *200*, 108182.
19. Fatai, A.; Jane, L.; Abdulazeez, A. Ensemble machine learning: An untapped modeling paradigm for petroleum reservoir characterization. *J. Petr. Sci. Eng.* **2017**, *151*, 480–487.
20. Liu, J.; Liu, J. Integrating deep learning and logging data analytics for lithofacies classification and 3D modeling of tight sandstone reservoirs. *Geosci. Front.* **2022**, *13*, 101311. [CrossRef]
21. Al-Mudhafar, W.; Abbas, M.; Wood, D. Performance evaluation of boosting machine learning algorithms for lithofacies classification in heterogeneous carbonate reservoirs. *Mar. Petrol. Geol.* **2022**, *145*, 105886. [CrossRef]
22. Onwuchekwa, C. Application of machine learning ideas to reservoir fluid properties estimation. In Proceedings of the SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 6–8 August 2018.
23. Hadavimoghaddam, F.; Ostadhassan, M.; Heidaryan, E.; Sadri, M.A.; Chapanova, I.; Popov, E.; Cheremisin, A.; Rafieepour, S. Prediction of dead oil viscosity: Machine learning vs. classical correlations. *Energies* **2021**, *14*, 930. [CrossRef]
24. Farzad, H.; Hamed, J.; Behzad, M.; Salahshoor, K.; Elkamel, A.A. A reinforcement learning approach for waterflooding optimization in petroleum reservoirs. *Eng. Appl. Artif. Intel.* **2019**, *77*, 98–116.
25. Zhang, K.; Wang, Z.; Chen, G.; Zhang, L.; Yang, Y.; Yao, C.; Wang, J.; Yao, J. Training effective deep reinforcement learning agents for real-time life-cycle production optimization. *J. Petrol. Sci. Eng.* **2022**, *208*, 109766. [CrossRef]

26. Priyanka, E.; Thangavel, E.; Gao, X.; Sivakumar, N. Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques. *J. Ind. Inf. Integr.* **2022**, *26*, 100272. [CrossRef]
27. Ambade, A.; Karnik, S.; Songchitruksa, P.; Sinha, R.R.; Gupta, S. Electrical Submersible Pump Prognostics and Health Monitoring Using Machine Learning and Natural Language Processing. In Proceedings of the SPE Middle East Intelligent Oil and Gas Symposium, SPE, Virtual, 18–19 October 2021; p. D011S004R003.
28. Kumar, A. Artificial neural network as a tool for reservoir characterization and its application in the petroleum engineering. In Proceedings of the Offshore Technology Conference, OTC, Houston, TX, USA, 30 April–3 May 2012; p. OTC-22967-MS.
29. Ahmadi, M.; Ebadi, M.; Shokrollahi, A.; Majidi, S.M.J. Evolving artificial neural network and imperialist competitive algorithm for prediction oil flow rate of the reservoir. *Appl. Soft. Comput.* **2013**, *13*, 1085–1098. [CrossRef]
30. Amirian, E.; Fedutenko, E.; Yang, C.; Chen, Z.; Nghiem, L. Artificial neural network modeling and forecasting of oil reservoir performance. In *Applications of Data Management and Analysis: Case Studies in Social Networks and Beyond*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 43–67.
31. Zhong, Z.; Carr, T. Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO_2–Reservoir oil minimum miscibility pressure prediction. *Fuel* **2016**, *184*, 590–603. [CrossRef]
32. Akande, K.; Owolabi, T.; Olatunji, S.; AbdulRaheem, A. A hybrid particle swarm optimization and support vector regression model for modelling permeability prediction of hydrocarbon reservoir. *J. Petrol. Sci. Eng.* **2017**, *150*, 43–53. [CrossRef]
33. Moosavi, N.; Bagheri, M.; Nabi-Bidhendi, M.; Heidari, R. Porosity prediction using Fuzzy SVR and FCM SVR from well logs of an oil field in south of Iran. *Acta Geophys.* **2023**, *71*, 769–782. [CrossRef]
34. Ahmadi, M.; Chen, Z. Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs. *Petroleum* **2019**, *5*, 271–284. [CrossRef]
35. Shafiei, A.; Tatar, A.; Rayhani, M.; Kairat, M.; Askarova, I. Artificial neural network, support vector machine, decision tree, random forest, and committee machine intelligent system help to improve performance prediction of low salinity water injection in carbonate oil reservoirs. *J. Petrol. Sci. Eng.* **2022**, *219*, 111046. [CrossRef]
36. Zhang, H.; Wu, W.; Wu, H. TOC prediction using a gradient boosting decision tree method: A case study of shale reservoirs in Qinshui Basin. *Geo. Sci. Eng.* **2023**, *221*, 111271. [CrossRef]
37. Yu, Z.; Wang, Z.; Jiang, Q.; Wang, J.; Zheng, J.; Zhang, T. Analysis of factors of productivity of tight conglomerate reservoirs based on random forest algorithm. *ACS Omega* **2022**, *7*, 20390–20404. [CrossRef]
38. Aslam, N.; Khan, I.; Alansari, A.; Alrammah, M.; Alghwairy, A.; Alqahtani, R.; Alqahtani, R.; Almushikes, M.; AL Hashim, M. Anomaly Detection Using Explainable Random Forest for the Prediction of Undesirable Events in Oil Wells. *Appl. Comp. Intell. Soft Comp.* **2022**, *2022*, 14. [CrossRef]
39. Hidayat, F.; Astsauri, T. Applied random forest for parameter sensitivity of low salinity water Injection (LSWI) implementation on carbonate reservoir. *Alex. Eng. J.* **2022**, *61*, 2408–2417. [CrossRef]
40. Adankon, M.; Cheriet, M. Model selection for the LS-SVM. Application to handwriting recognition. *Pattern. Recogn* **2009**, *42*, 3264–3270. [CrossRef]
41. Vapnik, V. Principles of risk minimization for learning theory. *Adv. Neural Inf. Proc. Syst.* **1991**, *4*, 831–838.
42. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.