energies

MDPI

Review

Review of Machine Learning Methods for Steady State Capacity and Transient Production Forecasting in Oil and Gas Reservoir

Dongyan Fan [1,2], Sicen Lai [2], Hai Sun [2,*], Yuqing Yang [3], Can Yang [2], Nianyang Fan [2] and Minhui Wang [2]

[1] State Key Laboratory of Deep Oil and Gas, China University of Petroleum, Qingdao 266580, China
[2] School of Petroleum Engineering, China University of Petroleum, Qingdao 266580, China
[3] Institute of Geology, Tenth Oil Production Plant, Daqing Oilfield Co., Ltd., Daqing 163000, China
* Correspondence: sunhai@upc.edu.cn

Abstract: Accurate oil and gas production forecasting is essential for optimizing field development and operational efficiency. Steady-state capacity prediction models based on machine learning techniques, such as Linear Regression, Support Vector Machines, Random Forest, and Extreme Gradient Boosting, effectively address complex nonlinear relationships through feature selection, hyperparameter tuning, and hybrid integration, achieving high accuracy and reliability. These models maintain relative errors within acceptable limits, offering robust support for reservoir management. Recent advancements in spatiotemporal modeling, Physics-Informed Neural Networks (PINNs), and agent-based modeling have further enhanced transient production forecasting. Spatiotemporal models capture temporal dependencies and spatial correlations, while PINN integrates physical laws into neural networks, improving interpretability and robustness, particularly for sparse or noisy data. Agent-based modeling complements these techniques by combining measured data with numerical simulations to deliver real-time, high-precision predictions of complex reservoir dynamics. Despite challenges in computational scalability, data sensitivity, and generalization across diverse reservoirs, future developments, including multi-source data integration, lightweight architectures, and real-time predictive capabilities, can further improve production forecasting, addressing the complexities of oil and gas production while supporting sustainable resource management and global energy security.

Keywords: machine learning; steady-state capacity prediction; transient production forecasting; spatio-temporal analysis; physics-informed neural networks

check for updates

Academic Editors: Konstantinos G. Kalogiannis and Maria Antoniadou

Received: 30 December 2024
Revised: 2 February 2025
Accepted: 6 February 2025
Published: 11 February 2025

Citation: Fan, D.; Lai, S.; Sun, H.; Yang, Y.; Yang, C.; Fan, N.; Wang, M. Review of Machine Learning Methods for Steady State Capacity and Transient Production Forecasting in Oil and Gas Reservoir. *Energies* **2025**, *18*, 842. https://doi.org/10.3390/en18040842

Copyright: © 2025 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/).

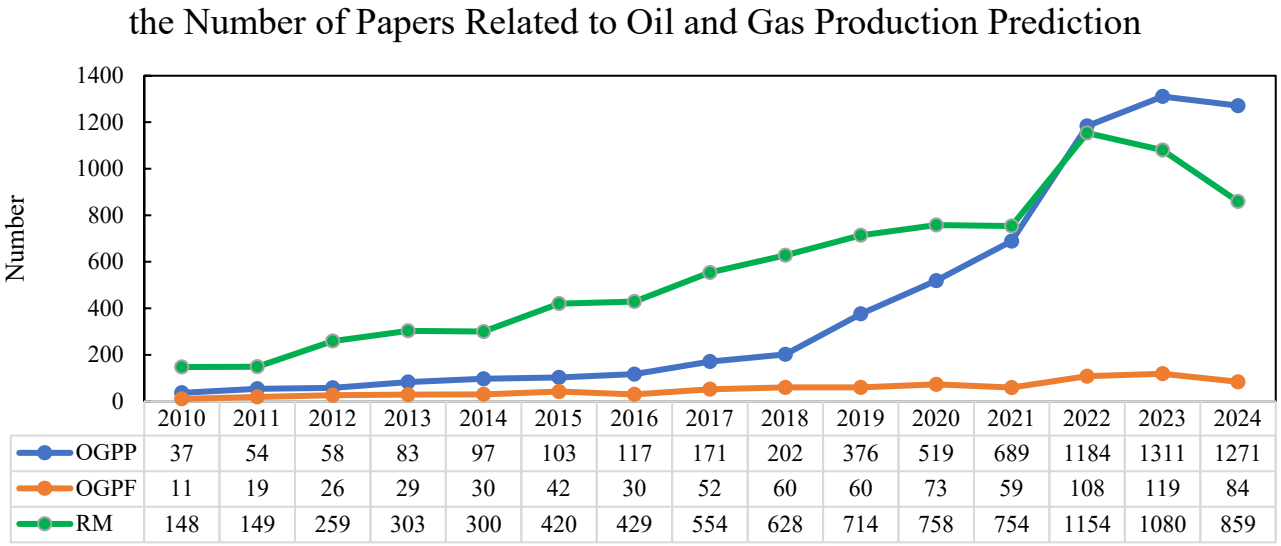
1. Introduction

Oil and natural gas, as pivotal energy resources, play indispensable roles in driving global economic growth and societal advancement [1]. Despite the abundance of global oil and gas reserves, significant challenges persist in oil and gas extraction, reservoir reserve estimation, and the optimization of transient production strategies, especially as traditional fields approach depletion and newly discovered reservoirs exhibit increasing complexity [2,3]. Ensuring a stable and sustainable energy supply necessitates accurate production forecasting, comprehensive steady state capacity assessments, and the development of efficient production strategies. These demands are not only critical for optimizing resource utilization but also essential for long-term energy security and sustainable development [4–6].

Steady state capacity and transient production forecasting, as a foundational technology in oil and gas field development, hold immense theoretical and practical significance [7]. Accurate forecasting enables the formulation of optimized development strate-

gies, enhances operational efficiency, and maximizes resource utilization [8,9]. Traditional methods, such as decline curve analysis (DCA) [9–12] and reservoir numerical simulations (RNS) [13–15], have been widely applied and provide valuable insights. However, as the focus shifts toward unconventional reservoirs, the limitations of these conventional approaches, including reduced accuracy and scalability, have become increasingly evident [16].

A bibliometric analysis of research publications from 2010 to 2024 reveals a significant increase in studies related to oil and gas production prediction (OGPP), reflecting the growing importance of this field. As shown in Figure 1, the number of publications in this area has grown steadily, from 37 in 2010 to 1311 in 2023. Research in oil and gas productivity forecasting (OGPF) has also increased, though at a slower pace, with annual publications rising from 11 in 2010 to 119 in 2023. Additionally, studies in reservoir modeling (RM) have exhibited notable growth, starting from 148 publications in 2010, peaking at 1154 in 2022, and slightly declining to 1080 in 2023. This trend underscores the increasing reliance on advanced computational and data-driven techniques to enhance production prediction accuracy and efficiency.



	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
OGPP	37	54	58	83	97	103	117	171	202	376	519	689	1184	1311	1271
OGPF	11	19	26	29	30	42	30	52	60	60	73	59	108	119	84
RM	148	149	259	303	300	420	429	554	628	714	758	754	1154	1080	859

Figure 1. Trends in the Number of Papers Related to Oil and Gas Production Prediction.

Recent advancements in computational power and algorithm development have catalyzed the adoption of machine learning (ML) techniques in production forecasting [17]. Machine learning is primarily categorized into supervised learning and unsupervised learning. Supervised learning involves training models using labeled data, where the algorithm learns the relationship between input features and known output values. Examples include regression models (e.g., linear regression, support vector regression) and classification models (e.g., decision trees, neural networks). The advantage of supervised learning is its high accuracy when sufficient labeled data are available; however, it heavily depends on data quality and requires significant manual labeling effort.

Unsupervised learning, on the other hand, is used when labeled data are unavailable. It identifies patterns and structures in data through clustering (e.g., K-means, DBSCAN) or dimensionality reduction (e.g., PCA, autoencoders). The advantage of unsupervised learning is its ability to discover hidden patterns without requiring labeled data, but it often produces results that are harder to interpret and evaluate.

Steady-state capacity forecasting and transient production forecasting typically fall under unsupervised learning, as they rely on recognizing historical data patterns without explicit input-output mappings. However, hybrid approaches that integrate supervised and unsupervised learning, such as semi-supervised and self-supervised learning, are emerging to enhance predictive accuracy and robustness [18]. ML algorithms, characterized by their ability to extract latent patterns from data, have demonstrated unparalleled potential in improving prediction accuracy and computational efficiency. Emerging methods such as deep learning frameworks, including recurrent neural network (RNN), convolutional neural network (CNN), and physics-informed neural network (PINN), alongside proxy models, provide innovative solutions to complex forecasting challenges. These techniques, tailored to address nonlinear dynamics, spatiotemporal dependencies, and intricate physical constraints, offer a scientific and reliable foundation for decision-making in oil and gas field development.

This paper aims to provide a comprehensive review of the application and optimization of ML techniques in steady-state capacity prediction and transient production forecasting. By systematically analyzing the strengths, limitations, and optimization strategies of various approaches, this study identifies current trends and future directions in ML-driven forecasting. Special emphasis is placed on integrating ML with physical constraints to enhance both prediction accuracy and interpretability. This paper is divided into time-independent steady-state capacity forecasting and time-dependent transient production forecasting. The transient production forecasting is further categorized into time series models, spatiotemporal models, physics-based models, and proxy models. The paper summarizes the latest forecasting methods and discusses future research directions. The paper is structured as follows: Section 2 reviews machine learning methods for steady-state capacity prediction. Section 3 examines transient production forecasting with a focus on spatiotemporal approaches. Section 4 discusses integrating physical constraints into predictions. Section 5 highlights proxy models for large-scale simulations. Section 6 summarizes challenges and future research directions. Lastly, the conclusions of this paper are provided in Section 7.

2. Steady-State Capacity Forecasting

Steady-state capacity forecasting is a fundamental aspect of oil and gas field development. Accurate assessment of reservoir productivity before production testing is critical for improving testing efficiency and informing development strategies [18]. In recent years, the rapid advancement of machine learning algorithms has led to the widespread adoption of both traditional machine learning and neural network methods in productivity prediction, yielding promising results [19]. Traditional machine learning approaches include linear regression (LR), support vector machines (SVM) [20–22], random forests (RF) [23–26], and gradient boosting (GB) [27–29]. Researchers have made significant progress in exploring these methods for capacity prediction. Figure 2 illustrates the workflow for productivity prediction, encompassing feature selection, data splitting, model training, nested optimization algorithms, and model evaluation. Common feature selection techniques include the gray correlation approach (GRA) [30], principal component analysis (PCA) [31], and K-means clustering [32].

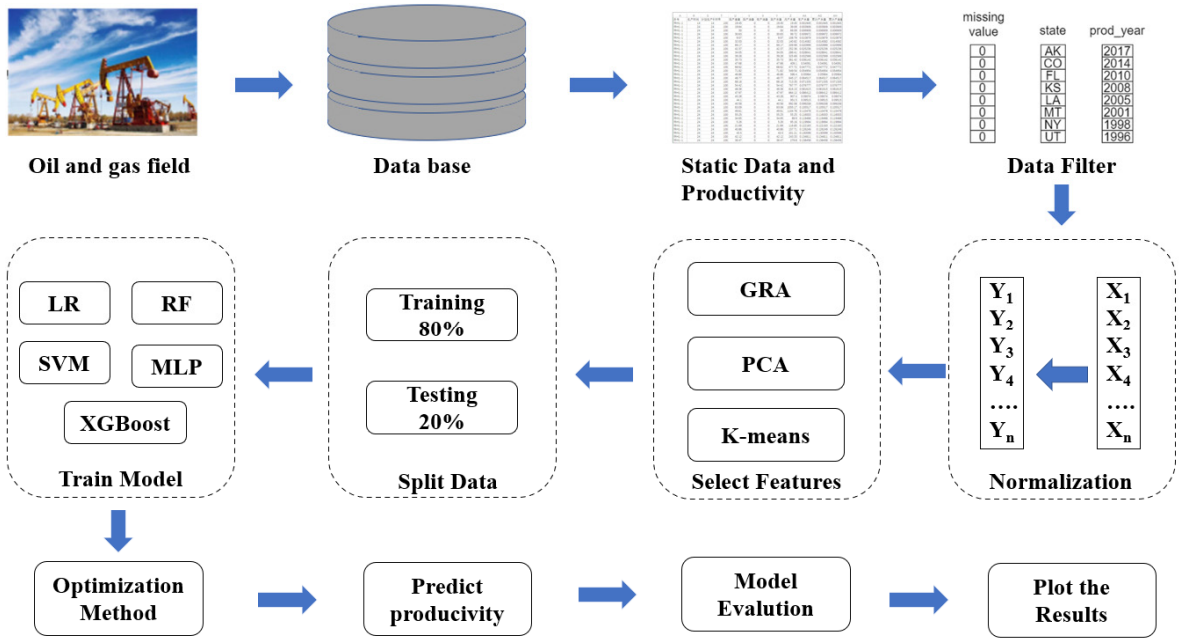


Figure 2. Workflow for productivity prediction.

2.1. Linear Regression Algorithm

Linear regression assumes a linear relationship between the input features and the target variable, meaning the output can be expressed as a weighted sum of the inputs plus a bias term. It also assumes that the independent variables are not highly correlated (no multicollinearity), the residual errors are normally distributed with a mean of zero, and homoscedasticity is present, meaning the variance of errors remains constant across all values of the independent variables [33].

Steady-state capacity in oil and gas fields is influenced by numerous factors, many of which are difficult to quantify. The relationship between these factors and productivity often exhibits complex nonlinear behavior [34]. Changes in individual factors may lead to nonlinear effects on productivity, while interactions among multiple factors can further amplify these effects. Consequently, accurate capacity forecasting requires sophisticated and advanced refined modeling techniques.

To address this challenge, researchers have proposed combining pre-regression techniques with regularization methods such as the Least Absolute Shrinkage and Selection Operator (Lasso) [35] and ridge regression (Ridge) [36]. Initially, pre-regression methods identify variables significantly associated with steady-state capacity, simplifying model structure and enhancing computational efficiency. Subsequently, Lasso and Ridge are applied to build the productivity prediction model. Experimental results demonstrate that this approach achieves high prediction accuracy on both training and test datasets, with relative errors controlled within 15%. These findings validate the effectiveness and accuracy of the strategy, offering robust technical support for oil and gas field development [35].

The model evaluation metrics used to assess the prediction results include the Mean Absolute Error (*MAE*), Mean Squared Error (*MSE*), Root Mean Squared Error (*RMSE*), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (*MAPE*). The calculation formulas are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|q_i - \hat{q}_i| \quad (1)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(q_i - \hat{q}_i)^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(q_i - \hat{q}_i)^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(q_i - \hat{q}_i)^2}{\sum_{i=1}^{n}(q_i - \bar{q}_i)^2} \quad (4)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{q_i - \hat{q}_i}{q_i}\right| \times 100\% \quad (5)$$

In the formula, q_i represents the actual value, $\bar{q}_i$ is the mean of all actual values, and $\hat{q}_i$ represents predicted value

Due to the nonlinear relationships between the influencing factors, simple linear regression may not capture these complexities, thus more advanced modeling techniques are required. By combining pre-regression methods with linear regression (LR), the model structure can be simplified, computational efficiency improved, and high prediction accuracy maintained.

2.2. Support Vector Machine

Support vector machines (SVM) are extensively utilized in traditional machine learning for regression and classification tasks, including steady-state capacity prediction of oil and gas wells, well classification, and fluid flow stage prediction [22,37]. SVM assumes that the data can be separated by a hyperplane, or in the case of nonlinear SVM, that a suitable kernel function can transform the data into a higher-dimensional space where it becomes linearly separable [38]. It also assumes a trade-off between maximizing the margin and minimizing classification errors, and the choice of kernel function significantly impacts performance.

Ibrahim demonstrated the effectiveness of SVM in predicting the steady-state capacity of a single well [39]. Similarly, Jessica achieved a prediction accuracy of 84% by combining stochastic gradient descent (SGD) with SVM [40]. highlighting the crucial role of optimization techniques in improving prediction efficiency and accuracy.

To further enhance SVM's predictive performance, researchers have focused on data preprocessing, such as dimensionality reduction using PCA [41], K-means clustering for screening fracturing construction parameters [42], and integrating multiple methods to predict shale oil steady-state capacity [43]. Additionally, structural optimizations, such as tuning SVM hyperparameters with the Gray Wolf Optimization (GWO) algorithm [44], have yielded improved results.

In conclusion, SVM performs excellently in predicting the steady-state capacity of oil and gas wells, effectively handling nonlinear relationships and performing both classification and prediction tasks. The use of optimization techniques, such as SGD and hyperparameter tuning with GWO, significantly enhances the prediction accuracy of SVM. Additionally, data preprocessing methods like dimensionality reduction using PCA and K-means clustering have further improved SVM performance. The combination of these techniques makes SVM highly applicable and effective in steady-state capacity prediction in the oil and gas industry.

2.3. Integrated Classification Methods

Integrated learning methods, particularly random forest (RF) classifiers, have demonstrated outstanding performance in recent years [45]. By combining multiple decision trees in parallel, RF effectively mitigates overfitting and significantly enhances prediction

accuracy and robustness [24]. With strong generalization capabilities, RF is suitable for both classification and regression tasks, particularly in high-dimensional and complex feature scenarios. Random forest assumes that the individual decision trees in the ensemble are independent and that each tree captures different aspects of the data by randomly selecting subsets of features. It also assumes that the selected features provide meaningful information for the prediction task and that an ensemble of trees will generalize better than a single tree. Sufficient training data are necessary to ensure diverse and robust trees.

For instance, Shi applied RF models to predict proven oil reserves, incorporating multidimensional indicators such as oil resources, geological conditions, and extraction technologies [46]. This approach overcame the limitations of traditional models, providing accurate and reliable predictions for exploration decision-making. Further improvements have been achieved through hyperparameter optimization, such as refining RF's performance with the Particle Swarm Optimization (PSO) method [47].

In conclusion, integrated learning methods, particularly Random Forest (RF) classifiers, excel in prediction tasks. By combining multiple decision trees, RF effectively mitigates overfitting, enhances prediction accuracy and robustness, and is especially suitable for high-dimensional and complex datasets. RF assumes that each decision tree captures different aspects of the data, thereby enhancing the model's generalization ability. The performance of RF is further improved through hyperparameter optimization techniques, such as Particle Swarm Optimization (PSO).

2.4. Extreme Gradient Boosting Algorithm

Extreme Gradient Boosting (XGBoost), a gradient boosting-based integrated learning method, offers high efficiency, accuracy, and interpretability [48]. XGBoost assumes that weak learners (decision trees) can be improved iteratively by correcting the residual errors of previous trees in a boosting framework. It follows an additive model where each tree refines the predictions of the previous ones. The model assumes that the loss function is differentiable for gradient-based optimization and that the data are not excessively noisy, as boosting methods are sensitive to overfitting on noisy datasets. By iteratively training multiple weak classifiers and combining them into a strong classifier, XGBoost effectively handles large-scale datasets and improves prediction accuracy [49]. For instance, Liao employed XGBoost to predict productivity in low-permeability reservoirs, incorporating feature relationships, multi-modeling, and k-fold cross-validation to enhance accuracy and robustness [50]. Similarly, Liu applied gradient boosting algorithms like XGBoost, LightGBM, and Gradient Boosting Decision Tree (GBDT)to predict the steady-state capacity of shale oil horizontal wells, identifying key features that inform extraction strategies [51].

XGBoost is an efficient, accurate, and interpretable integrated learning method that iteratively optimizes weak learners (decision trees) by correcting the residual errors of previous trees, thereby improving prediction accuracy. XGBoost is suitable for large-scale datasets and effectively handles complex prediction tasks, especially in applications such as predicting productivity in low-permeability reservoirs and steady-state capacity in shale oil wells. It relies on differentiable loss functions for gradient-based optimization and is sensitive to noise, emphasizing the importance of clean and high-quality data. By combining multi-modeling, feature relationships, and cross-validation, XGBoost provides high accuracy and robustness in its predictions, demonstrating its strong potential in the oil and gas industry.

2.5. Neural Network Approach

Multilayer Perceptron (MLP), a type of artificial neural network (ANN), consists of multiple hidden, input, and output layers. MLP assumes that the input data can be transformed into a separable form through nonlinear activation functions. It also assumes that the training data are independent and identically distributed (IID), ensuring stable learning. The model relies on sufficient training data to learn complex patterns and assumes that hyperparameters, such as learning rate, network depth, and activation functions, are properly tuned to avoid issues like vanishing or exploding gradients. MLP excels at learning complex nonlinear relationships and is widely applied in productivity prediction [52]. However, MLP models are highly sensitive to hyperparameters, and incorrect configurations can lead to significant computational costs [53]. Table 1 outlines key factors affecting prediction accuracy and corresponding strategies for improvement.

Table 1. Factors Affecting Capacity Prediction Accuracy and Coping Strategies.

Factors	Influence on Prediction Accuracy	Coping Strategies
Data Quality	Noise, missing values, and outliers reduce model robustness	Improve data collection, cleaning, and imputation methods
Feature Selection and Dimensionality Reduction	Poor selection causes overfitting and reduces accuracy	Use ML/statistical methods and regularization
Model Complexity and Overfitting	Excessive complexity leads to overfitting	Apply regularization, cross-validation, and early stopping
Nonlinear Relationship	Unaddressed nonlinearities bias predictions	Employ models handling nonlinearity

Wang applied ANN methods to predict and analyze the steady-state capacity of the Bakken shale reservoir [54]. To improve robustness of ANN, k-fold cross-validation was implemented, while Sobol analysis was used to examine engineering parameters influencing production. Additional techniques, such as Kalman filtering for data preprocessing, further alleviated the impact of missing data, optimized computational speed, and enhanced prediction accuracy [55]. Jiang combined ANN with conservation of mass equations in a control volume model to indirectly predict production capacity changes, enhancing both prediction accuracy and interpretability [56].

MLP is a type of ANN that excels at learning complex nonlinear relationships and is widely used in productivity prediction. MLP assumes input data can be transformed into a separable form through nonlinear activation functions, and requires IID training data for stable learning. While MLP is effective at capturing intricate patterns, it is highly sensitive to hyperparameter settings, and incorrect configurations can lead to significant computational costs. To improve robustness and prediction accuracy, various strategies have been employed, such as k-fold CV (Cross-Validation), Sobol analysis, and KF (Kalman Filtering) for data preprocessing. Additionally, combining MLP with physical models, like CMM (Conservation of Mass equations), can enhance both accuracy and interpretability of predictions.

Different prediction models show varying accuracy and advantages in forecasting steady-state capacity for oil and gas field development. Table 2 summarizes the characteristics, applicable scenarios, advantages, and limitations of different methods for capacity prediction. Linear regression, due to its simplicity and the assumption of a linear relationship between input features and the target variable, typically provides lower prediction accuracy (R^2 around 0.5–0.7) and struggles with complex, nonlinear data. Support Vector Machines (SVM) offer moderate accuracy (R^2 around 0.6–0.8) for small to medium datasets, but their performance is highly dependent on the choice of kernel function and can be

slow for large datasets. Random Forest performs relatively well with higher accuracy (R^2 around 0.7–0.9) due to its ability to handle nonlinear relationships, though it comes with higher computational cost. XGBoost, with its strong ensemble learning capability, often achieves higher accuracy (R^2 around 0.8–0.95), but its sensitivity to noisy data can reduce its robustness. Multi-layer Perceptrons (MLP) can capture complex nonlinear patterns with accuracy ranging from 0.8 to 0.95, but require large datasets and substantial computational resources, and are sensitive to hyperparameters [57–59].

Table 2. Applicability and Pros/Cons of Capacity Prediction Methods.

Algorithm Type	Characteristics	Applicable Scenarios	Advantages	Limitations
Linear Regression Method	Simple, interpretable	Typically < 1000 samples	Low computation, easy to explain	Poor for complex, nonlinear data
SVM	Flexible, handles complex boundaries	100–5000 samples	Works well with small samples	Slow for large datasets
Integrated Classification Methods	Combines multiple models for accuracy	500–50,000 samples	Reduces overfitting, handles high dimensions	High computation, long training time
MLP	Learns complex nonlinear patterns	10,000+ samples	Powerful learning, adapts to complex tasks	Hard to tune, resource-intensive

Optimization algorithms significantly improve the accuracy of prediction models by fine-tuning hyperparameters. For example, regularization and cross-validation help linear regression reduce overfitting and improve accuracy. For SVM, optimizing kernel functions and penalty parameters enhances its performance, improving R^2 to 0.75–0.85 [60]. Random Forest and XGBoost, by adjusting tree depth, learning rates, and regularization terms, can achieve R^2 values between 0.85 and 0.98, improving their ability to handle complex nonlinear data [61,62]. MLP can improve its accuracy by fine-tuning hidden layers and learning rates, helping to reduce overfitting [63]. Overall, optimization algorithms refine hyperparameter settings, boosting model generalization and accuracy, making them more effective for handling complex, high-dimensional data in oil and gas field development.

3. Considering Spatiotemporal Sequence Production Prediction

Transient production is a critical indicator of oil and gas field development, and accurate prediction of future production is essential for analyzing reservoir development dynamics [64]. Changes in single-well production are influenced by various factors such as reservoir properties and production enhancement measures. To achieve accurate single-well production prediction, it is necessary to comprehensively consider the temporal and spatial series of production and understand the characteristics of production changes. Conventional production forecasting methods include production DCA, RNS and analytical methods [65]. However, machine learning methods, with their stronger nonlinear fitting ability and higher efficiency, show greater potential, particularly for transient production prediction [66].

Building on single models, researchers have optimized predictions through hyper-parameter selection, data preprocessing and elimination of outliers to make model more accurate [67]. Figure 3 demonstrates the workflow for transient production prediction considering spatio-temporal series, involving feature selection, dataset partitioning, model selection, model training, and evaluation.

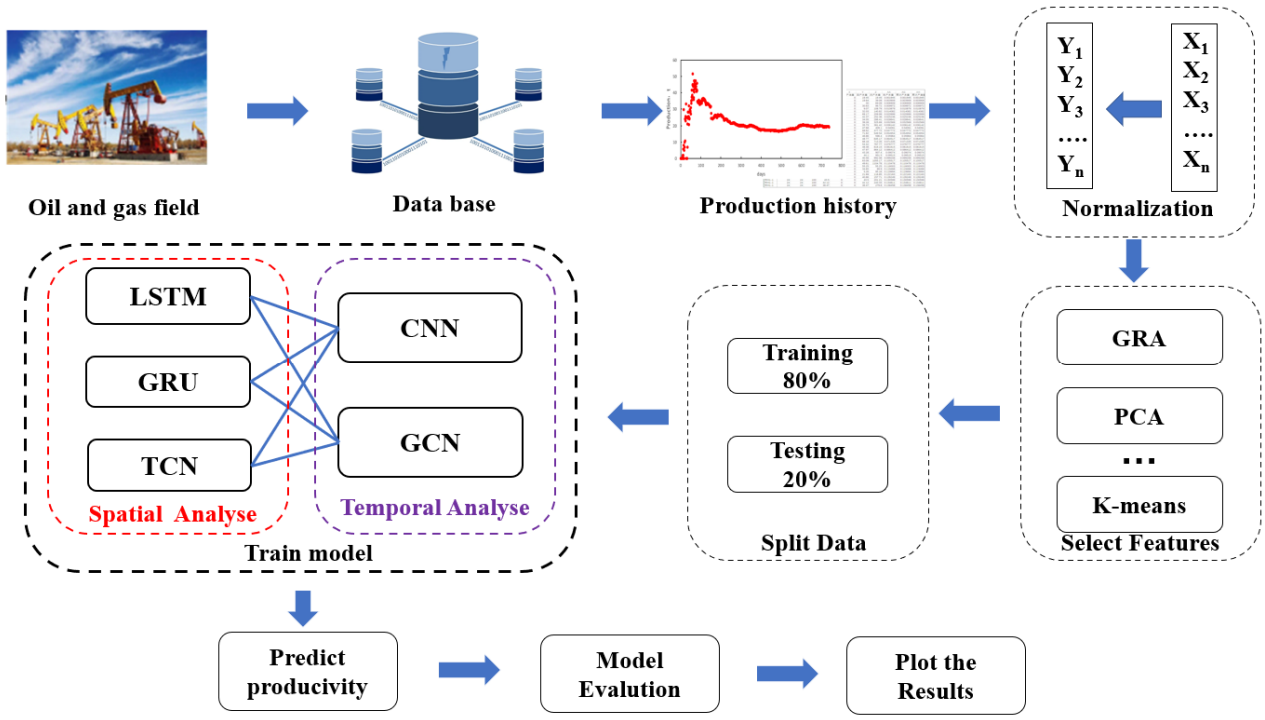


Figure 3. Spatio-temporal series yield prediction process.

3.1. Considering Time Series Transient Production Prediction

Numerous machine learning algorithms address production forecasting, with prominent models including MLP, RNN, and CNN. MLP updates parameters via backpropagation to predict production [68], RNN effectively handles temporal dependencies through its recurrent structure [69], and CNN excels in feature extraction and spatial hierarchy recognition [70]. However, due to inherent limitations in these models, Table 3 shows researchers have optimized their structures to better align with the characteristics of oil and gas production data, resulting in improved accuracy and broader applicability in production prediction tasks.

Table 3. Methods for Improving Transient Production Prediction Accuracy.

Optimization Method	Advantages	Challenges and Limitations
Hyperparameter Optimization	Significantly improves model accuracy, reduces overfitting	Complex process, computationally intensive, time-consuming
Data Preprocessing	Reduces noise, improves data stability, enhances training	Requires deep understanding of data, risks information loss
Feature Selection	Identifies patterns using clustering, optimizes inputs	Results depend on initial clustering centers, may be unstable
Ensemble Techniques	Handles linear and nonlinear data, improves accuracy	Requires careful model/algorithm selection, risk of overfitting or heavy computational load

3.1.1. Multi-Layer Perceptron

MLP is a foundational model for time series prediction problems. Tian developed a back-propagation (BP) neural network model using time, cumulative production, and formation pressure as inputs and shale gas daily production as output [71]. To enhance model performance, a genetic algorithm (GA) was employed to optimize BP network bias

and weights, significantly improving prediction accuracy. Zhou proposed a combined BP-LSTM model, further improving prediction accuracy by integrating empirical constraints with the BP-LSTM algorithm [72].

3.1.2. Recurrent Neural Network and Variants

RNN, particularly Long Short-Term Memory (LSTM) networks, have been widely applied in production prediction. LSTM addresses long-term dependencies through gating mechanisms. Traditional RNNs assume that short-term dependencies in time series can be captured through recursive structures, but they may struggle with long-term dependencies, leading to issues like vanishing or exploding gradients [73]. LSTM and Gated Recurrent Unit (GRU) are improvements that address these problems and enhance the model's ability to handle long-term dependencies. Hou demonstrated the superiority of LSTM over traditional backpropagation (BP) methods by utilizing oilfield production data spanning 13 years and incorporating Dropout layers to prevent overfitting [74].

Researchers have focused on hyperparameter optimization and data preprocessing to improve LSTM's prediction accuracy. For example, Song optimized LSTM parameters using a particle swarm optimization (PSO) algorithm, significantly enhancing production prediction accuracy in fractured horizontal wells [75]. Liu further improved LSTM input processing by combining integrated empirical modal decomposition (EEMD) with dynamic time regularization (DTW), achieving higher prediction precision [76].

Some scholars have combined multiple LSTM models through integration methods, improving accuracy and robustness [77]. Fan divided production forecasts into linear and residual parts, using ARIMA for linear trends and LSTM for residual variations [78]. This hybrid approach enhanced prediction accuracy and provided a comprehensive understanding of yield change patterns.

However, LSTM suffers from a large number of parameters and low convergence rates, especially when handling large datasets, resulting in poor processing capabilities. To address these issues, Cho proposed the GRU based on LSTM [79]. GRU simplifies the internal structure of LSTM, reduces network training time, and maintains a certain level of prediction accuracy. Li introduced a novel framework combining a bidirectional gated recurrent Unit (Bi-GRU) and the sparrow search algorithm (SSA) [80]. The Bi-GRU fully exploits past and future information from production sequences and relevant features, while SSA is utilized for hyperparameter tuning of the Bi-GRU model, thereby enhancing the accuracy of oil production rate prediction. Gao developed GRU and LSTM-based yield prediction models, integrating regression interpolation for noise reduction and differencing methods for data preprocessing [81]. These approaches address issues such as outliers and periodic fluctuations in oilfield production data, ultimately improving model prediction accuracy. Li proposed a combined neural network model based on Bi-GRU and Multi-Task Learning (MTL) [82]. MTL integrates multi-stage production forecasting tasks by sharing task-related representations to enhance prediction accuracy. It also adjusts the loss function weights to improve the model's fitting ability. Du structurally improved the GRU to extract external economic information and incorporated autoregressive models to embed prior knowledge, enhancing the model's capability to capture temporal sequence information [83]. This improvement not only increases prediction accuracy but also accounts for the impact of economic changes on production variations.

In recent years, RNN and their variants, particularly LSTM networks, have been widely applied in oilfield production forecasting. By incorporating techniques such as dropout layers and transfer learning, LSTM has demonstrated superior performance in prediction tasks. To further enhance prediction accuracy, researchers have explored areas such as

hyperparameter optimization, data preprocessing, and model integration, employing techniques like PSO and integrating EMD with GA [84].

Additionally, the GRU, a variant of LSTM, simplifies the internal structure to reduce training time, while the Bi-GRU leverages both past and future information in sequences for improved predictions. Figure 4 illustrates the network structures of RNN, LSTM, and GRU. The integration of methods such as regression interpolation, differencing, and MTL has further enhanced the accuracy and robustness of predictions. Moreover, structurally improved GRUs can extract external economic information and incorporate autoregressive models to enhance the ability to capture temporal sequence information, thereby better accounting for the impact of economic changes on production variations.

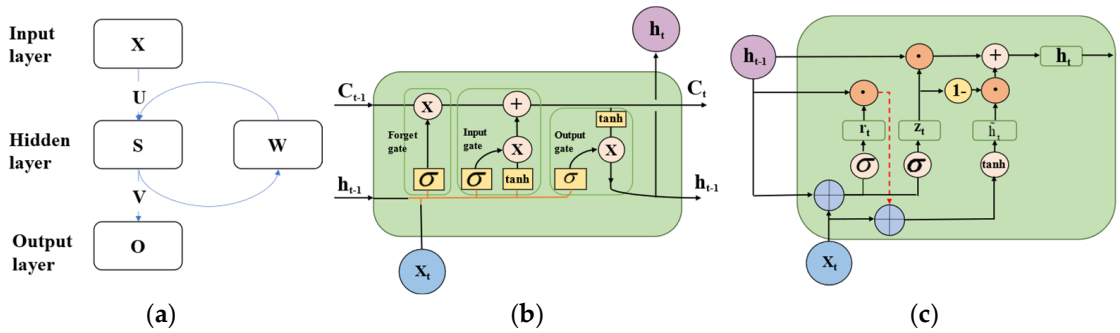


Figure 4. Machine learning models for time-series analysis: (**a**) RNN (**b**) LSTM (**c**) GRU.

RNN and its variants, especially LSTM, have demonstrated significant effectiveness in production forecasting by addressing long-term dependencies in time series data. Optimizations such as hyperparameter tuning, advanced preprocessing, and hybrid integrations have further improved accuracy and robustness. Additionally, simplified variants like GRU and frameworks like Bi-GRU optimized with algorithms such as SSA have reduced computational demands while maintaining accuracy. Despite these advancements, challenges in scalability, computational efficiency, and overfitting remain, highlighting the need for lightweight architectures, automated tuning, and advanced noise mitigation techniques to enhance future production prediction applications.

3.1.3. Convolutional Neural Networks and Variants

Convolutional Neural Networks assumes that the data have local features that can be captured through convolution operations. The convolution operation assumes that the input data exhibits translation invariance in space or time, meaning that the same feature will behave similarly at different locations [84]. CNN extracts higher-order features of the data layer by layer, making it effective for capturing local patterns, particularly in image and time-series data. Originally designed for image processing, CNNs have been adapted to time-series data by incorporating one-dimensional convolutions and structural modifications. Temporal Convolutional Networks (TCNs), for instance, introduce causal convolutions and residual concatenations, enabling the model to handle variable-length sequences while avoiding gradient vanishing and explosion issues in RNNs [85]. Additionally, adjustable convolution kernel sizes enhance adaptability to varying time series lengths, further improving prediction accuracy [86,87]. Figure 5 illustrates the structures of CNN and TCN models structure.

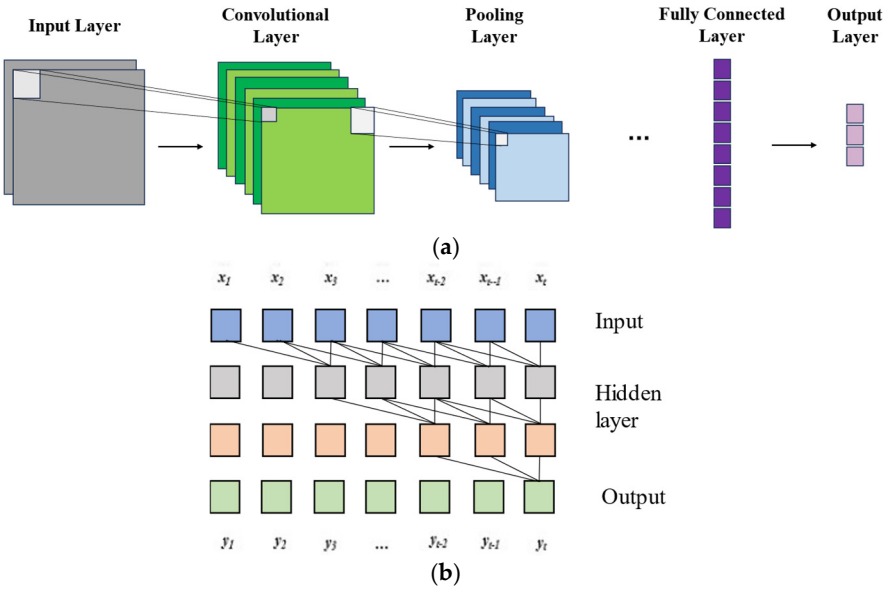


Figure 5. Machine learning models for temporal order analysis: (**a**) CNN (**b**) TCN.

To improve CNN's prediction accuracy, Zheng proposed a TCN model combined with an attention mechanism [88]. By expanding TCN's sensory field through causal convolution, the model processes longer time-series information with fewer network layers. Residual connectivity enhances information transfer across layers, addressing network degradation issues. The attention mechanism focuses on critical input sequence components, further boosting prediction accuracy. Similarly, Zhang optimized TCN hyperparameters using SSA, enhancing performance and generalization capabilities [89]. SSA (Sparrow Search Algorithm) is a population-based optimization algorithm that effectively explores the solution space to identify the optimal solution.

3.1.4. Multi-Model Production Prediction

Building on CNN's ability to extract features, many scholars have combined CNN with other models to achieve higher prediction accuracy [53]. Zha proposed a CNN-LSTM model for gas well production prediction, where CNN extracts features and LSTM captures sequence-dependent learning [90]. This combination proved more effective than standalone CNN or LSTM models in describing production trends. Song developed a horizontal well production prediction model using a hybrid neural network (CNN-LSTM), which incorporates both production and engineering factors as inputs, improving the model's accuracy [91].

Researchers have also refined 'CNN+' models based on prior studies. Li extracted temporal features with CNN and fed them into an LSTM model optimized by the PSO algorithm [92]. This approach significantly enhanced yield prediction accuracy. Zhou [93] introduced the Duong-CNN-LSTM (D-C-L) model, which integrates CNN and LSTM with the Duong decline model. Residuals from the Duong method guided further iterations of the CNN-LSTM model, resulting in higher accuracy, efficiency, and cost-effectiveness for shale oil production prediction compared to conventional methods.

Similarly, Chen developed a CNN-GRU model for enhanced oil recovery (EOR) prediction [93]. CNN layers extracted relevant features affecting oil production, while GRU layers modeled temporal relationships. Bayesian optimization was used to fine-tune the model's hyperparameters, achieving robust prediction performance.

Various machine learning models, including TCN, GRU, LSTM, and MLP, exhibit different levels of effectiveness in spatiotemporal sequence production prediction. Table 4

summarizes the prediction accuracy, recommended dataset length, and the main advantages and disadvantages of each model [73,84].

Table 4. Summary of Basic Machine Learning Models for Production Prediction.

Model	Prediction Accuracy Range	Recommended Dataset Length (Days)	Advantages	Disadvantages
MLP	75–80%	200–1000	Simple, fast convergence	Overfitting, limited sequence learning
LSTM	82–88%	1000–5000	Captures long-term dependencies	High computation, slow training
GRU	85–89%	1000–5000	Faster training, fewer parameters	Less expressive than LSTM
TCN	87–91%	5000–10,000	Efficient time-series modeling	Requires fine-tuning

In summary, integrating MLP, CNN, and RNN's advantages with multi-model optimization strategies provides an effective solution for oil and gas well production prediction. Future research should explore novel architectural combinations and optimization techniques, particularly for large-scale and complex datasets, to further improve accuracy and broaden applications to other time series prediction tasks.

3.2. Considering Temporal-Spatial Sequence Production Prediction

While temporal dependencies play a significant role in yield prediction, the spatial context, such as well geography and reservoir heterogeneity, also critically impacts production. Many researchers have integrated CNN's feature extraction capabilities with other models to incorporate spatial sequence effects on production [94].

Fargalla proposed TimeNet, a hybrid model combining CNN, Bi-directional Gated Recursive Units (BiGRU), Attention Mechanisms (AM), and Time2Vec [95]. The Time2Vec module effectively captures complex temporal patterns, reducing preprocessing complexity, while BiGRU identifies advanced temporal features and trends. Attention mechanisms enhance prediction accuracy by focusing on critical embedded information, and these integrated components collectively improve learning outcomes and overall prediction precision.

Pan introduced the CNN-LSTM-SA model, which integrates CNN, LSTM, and a self-attention mechanism (SA) to explore hidden spatiotemporal correlations in well production data [96]. The self-attention mechanism captures internal dependencies with high accuracy, leading to more robust and reliable predictions. Similarly, Mandella combined BiGRU and AM within a CNN framework to extract spatiotemporal features, achieving significant improvements in predictive accuracy and reliability [97]. The BiGRU network can provide new, more detailed temporal information or capture irregular trends in time series components. Additionally, the Attention Mechanism (AM) aids in understanding the impact of intrinsic information on ensuring learning accuracy.

Another model that can consider spatiotemporal sequences is the coupling model based on Graph Convolutional Networks (GCN). GCN assumes that the data can be modeled using a graph structure, where the relationships between nodes are represented by an adjacency matrix, and node features are propagated through graph convolution [98]. GCN assumes that node features should smooth out across neighboring nodes, meaning that features of adjacent nodes should be similar. The graph convolution operation helps the model capture the structural information between nodes, making it especially suitable for non-Euclidean data such as social networks or molecular structures. Du developed a deep learning framework that integrates Graph Convolutional Networks (GCNs) with GRU [99]. This framework includes an adaptive correction module using GCN to refine inter-well connectivity based on dynamic production data and a GRU-based temporal learning

module for well performance prediction. By addressing reservoir-specific conditions and inter-well connectivity, this approach effectively resolves challenges in dynamic production forecasting. Many researchers have explored the use of Graph Convolutional Networks (GCN) to address spatial sequence problems, incorporating features from neighboring wells into production forecasting models [100].

Further extending GCN applications, Du introduced the Multi-Graph Convolutional Networks (MGCN) framework [101]. This model captures well interactions from various interrelationship maps, including well patterns, spatial distances, and injection-extraction dynamics. The extracted features are processed using LSTM and fully connected networks to fuse spatiotemporal correlations with production factors, culminating in highly accurate predictions. The empirical mode decomposition (EEMD) component threshold denoising method is applied to preprocess real oilfield data. Bayesian optimization fine-tunes hyperparameters to achieve optimal performance. However, GCN-based models often require extensive training datasets and exhibit limited generalization capabilities when encountering previously unseen nodes.

To address these challenges, Xue adopted the Graph Sampling and Aggregation (GraphSAGE) architecture. This method improves generalization to unseen nodes by enabling mini-batch training on large graphs [102]. Coupled with GRU, the ST-GraphSAGE model captures temporal interactions between target wells and neighboring wells, effectively predicting future production. Similarly, Zhao proposed the T-DGCN model, which integrates temporal, spatial, and geologic features [103]. In this model, Dynamic Time Warping (DTW) measures geological similarities, while GRU extracts temporal features to forecast daily gas production with high precision. For practical applications, algorithm adaptability, multi-source data fusion, and real-time predictive capabilities will be critical research focuses. Breakthroughs in these areas will further elevate model performance.

Temporal-spatial sequence models have been increasingly applied to production prediction, addressing the dual challenges of temporal dependencies and spatial factors such as well geography and reservoir heterogeneity. Table 5 shows common methods for spatiotemporal sequence prediction. Advanced hybrid models, including TimeNet, CNN-LSTM-SA, and BiGRU-AM frameworks, effectively combine CNN, RNN variants, and attention mechanisms to enhance feature extraction and prediction accuracy. Graph-based models, such as GCN-GRU, MGCN, GraphSAGE-GRU, and T-DGCN, further integrate spatial correlations and temporal dynamics, leveraging advanced techniques like Dynamic Time Warping (DTW) and multi-graph architectures to achieve precise predictions. Despite these advancements, challenges persist, including high computational complexity, extensive data requirements, and limited generalization to unseen nodes. Future research should focus on developing adaptable algorithms, integrating multi-source data, and enabling real-time predictive capabilities to enhance model scalability, efficiency, and robustness in complex reservoir conditions.

Machine learning models for production prediction have significantly evolved to address the dual challenges of temporal dependencies and spatial heterogeneity in reservoir conditions. Techniques like MLP, RNN, LSTM, GRU, and CNN have been optimized through hyperparameter tuning, data preprocessing, and hybrid integrations, improving accuracy and robustness. Advanced spatiotemporal models, such as TimeNet, CNN-LSTM-SA, and BiGRU-AM, effectively combine feature extraction, temporal learning, and attention mechanisms to enhance prediction precision. Graph-based models, including GCN-GRU, MGCN, GraphSAGE-GRU, and T-DGCN, further incorporate spatial correlations and temporal dynamics using innovative methods like multi-graph architectures and Dynamic Time Warping (DTW). Despite these advancements, challenges like high computational costs, extensive data requirements, and generalization limitations persist.

Future efforts should prioritize algorithm adaptability, real-time predictive capabilities, and multi-source data integration to enhance scalability and reliability in complex oilfield production scenarios.

Table 5. Common Methods for Spatiotemporal Sequence Prediction.

Method Category	Specific Methods	Characteristics and Summary
Time Series Prediction Methods	MLP, DNN	Suitable for time series data, but weak at capturing long-term dependencies.
	RNN and Variants	Optimized for temporal dependencies, widely used in oilfield production prediction.
	CNN and Variants	Handles sequential data well with 1D convolutions, improving prediction accuracy.
	Multi-Model Coupling	Combines CNN, LSTM, GRU to improve time feature extraction and predictive capabilities.
Spatiotemporal Sequence Methods	GCN/CNN+GRU/LSTM	Extracts spatiotemporal correlations by combining spatial features (GCN) with temporal models (GRU/LSTM), improving prediction accuracy.

4. Considering Physical Constraint Production Prediction

In spatio-temporal sequence yield prediction, traditional data-driven models often lack physical interpretability [104]. Physics-Informed Neural Networks (PINNs) address this gap by embedding physical laws and domain knowledge into machine learning models [105]. PINNs integrate physical constraints as regularization terms in the neural network loss function, enhancing interpretability, generalization, and adaptability. By incorporating physics, PINNs effectively handle noisy data, reducing the negative impact of noise on model performance.

However, challenges of PINNs, such as computational scalability and data sparsity, are not adequately addressed [106]. These challenges become more prominent when dealing with large-scale reservoir simulations or complex models where high-dimensional spatio-temporal data are involved [107]. Potential solutions for overcoming these computational challenges include domain decomposition and parallel computing. Domain decomposition involves dividing the computational domain into smaller sub-domains, allowing the model to handle smaller chunks of data more effectively. Parallel computing enables simultaneous processing of these sub-domains, significantly reducing computation time and improving scalability [108].

In oil and gas production prediction, PINNs demonstrate significant advantages by combining physical laws with neural networks' learning capabilities. For instance, transforming the Arps production decline model into partial differential equations allows PINNs to accurately identify shale gas well decline curves. Techniques such as Caputo fractional-order derivatives, experimental data, automatic differentiation, and sparse regression further improve predictions [109]. Introducing non-chiral fractional-order differential equations enhances accuracy and physical validity.

PINNs also excel in solving complex fluid physics problems. Integrating physical information reduces prediction bias while improving interpretability and portability [110,111]. To address data sparsity in large-scale reservoir simulations, researchers developed domain decomposition-based PINN architectures (PINN-DD) [112]. These architectures divide computational domains into sub-domains, leveraging PINN's ability to process small-scale spatio-temporal data effectively.

In conclusion, PINNs' integration of physical laws with machine learning offers broad application prospects in oil and gas production prediction. Figure 6 highlights key factors affecting PINN performance. Continued research in this area promises more accurate, reliable, and interpretable production predictions, advancing reservoir dynamic studies and supporting future exploration and production strategies.

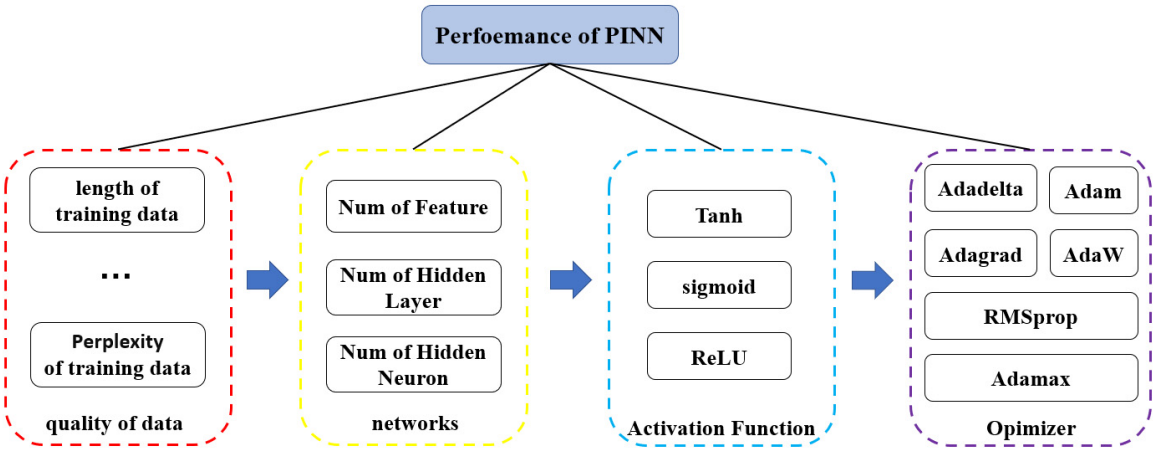


Figure 6. Key factors influencing PINN prediction performance.

5. Yield Prediction by Agent Modeling

Agent modeling, as an innovative tool, integrates measurable data with numerical simulation models to significantly enhance the reliability and accuracy of predictions while effectively reducing computational cost and time consumption [113]. This technique holds great potential in the development and utilization of oil and gas resources, particularly for production prediction. Traditional proxy models have contributed to improving prediction efficiency but face limitations in handling reservoir complexity and data requirements. With the rapid development of artificial intelligence and data mining technologies, intelligent agent models have emerged to address these challenges by accurately simulating dynamic reservoir properties and generating high-precision predictions within shorter timeframes. These intelligent agent models can accurately simulate the dynamic reservoir properties of an entire oilfield and generate high-precision prediction results within a short time. This capability optimizes decision support for oil and gas production, promoting more efficient and cost-effective development of oil and gas resources [114,115].

Machine learning plays a critical role in surrogate models by efficiently replacing traditional numerical simulations. By training deep learning models, such as convolutional neural networks (CNNs), surrogate models can rapidly learn the complex nonlinear relationships between reservoir parameters and production rates, enabling efficient prediction of reservoir states like pressure distribution and water saturation [87]. Compared to traditional reservoir numerical simulations, the advantages of surrogate models include faster computation, lower cost, and higher accuracy, allowing for real-time updates and optimized decision-making, making them well-suited for large-scale reservoir management and intelligent oilfield development [116]. The integration of machine learning with agent-based models further enhances their predictive capabilities, enabling more robust and adaptive forecasting methods for oil and gas production.

Proxy models play a vital role in oilfield production prediction, offering fast and accurate insights into the dynamics of complex systems. For instance, Dong applied migration learning with an LSTM model trained on historical data to predict production in wells with shorter development timelines [113]. Similarly, Rodriguez et al. evaluated agent

models for predicting oil and water production in water-injection fields by adjusting input parameters [116].

Studies have demonstrated the effectiveness of agent models in various scenarios. Mohaghigh et al. addressed data quality and coverage issues by coupling numerical simulations with agent models, ensuring robust predictions [117]. Khazaeni et al. employed 40 years of production data to predict both new well outputs and overall field performance, highlighting the high accuracy of agent models in data-rich environments [118]. Additionally, Aboaba improved prediction accuracy and robustness by developing a unified network of multiple neural networks [119,120].

Despite these advancements, agent models face challenges such as sensitivity to data quality and biases, especially with incomplete datasets. The necessity to combine agent models with numerical simulations for reliability adds to their complexity. Furthermore, constructing and optimizing proxy models demands significant computational resources and extensive data support, increasing their overall cost and complexity. However, continued advancements in machine learning techniques, particularly in deep learning and reinforcement learning, present opportunities to further optimize agent models, improving their adaptability and efficiency for oil and gas production forecasting.

6. Discussion

The application of machine learning (ML) and artificial intelligence (AI) in oil and gas production forecasting has provided significant improvements in accuracy, computational efficiency, and interpretability. The rapid evolution from traditional methods, such as decline curve analysis and reservoir numerical simulations (RNS), to advanced data-driven approaches highlights the potential of modern ML techniques. However, challenges including data sparsity, computational complexity, and the need for practical deployment of interpretable models remain critical obstacles. This section discusses key developments, ongoing challenges, and future directions in production prediction.

Steady-state capacity forecasting focuses on estimating a reservoir's productivity under stable production conditions. Traditional methods such as linear regression (LR) and decline curve analysis (DCA) often provide simple yet suboptimal results due to their limited ability to model nonlinear relationships. Advanced ML techniques, such as Support Vector Machines (SVM), Random Forests (RF), Extreme Gradient Boosting (XGBoost), and neural network models like Multi-Layer Perceptron (MLP), have significantly improved prediction performance.

The use of optimization techniques, including feature selection, regularization, and hyperparameter tuning (e.g., Particle Swarm Optimization and Bayesian optimization), has led to further performance gains, reducing overfitting and enhancing model robustness. However, these improvements are still constrained by the availability of high-quality training data and the need for scalable computing frameworks to handle large datasets efficiently.

Transient production forecasting is essential for analyzing oilfield dynamics under fluctuating production conditions. Time series-based models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs) have shown significant success in capturing production dynamics. Their ability to capture long-term dependencies makes them well-suited for transient production prediction.

Hybrid architectures, such as CNN-LSTM and BiGRU-AM, further enhance accuracy by combining spatial and temporal learning capabilities. Graph Convolutional Networks (GCN), integrated with temporal models like GRU, model inter-well spatial connectivity while capturing reservoir heterogeneity. However, one of the primary challenges remains the computational burden associated with training such models, particularly when integrating real-time streaming data from production wells. The development of more

computationally efficient architectures, such as sparse neural networks and knowledge distillation methods, could provide practical solutions to these issues.

Physics-Informed Neural Networks (PINNs) incorporate physical principles (e.g., Darcy's law, conservation of mass) as constraints in ML models, improving interpretability and reliability. By embedding physics-based knowledge into ML algorithms, PINNs address the shortcomings of purely data-driven models and improve generalizability in noisy and sparse datasets. PINNs have demonstrated particular success in modeling shale gas decline curves and fluid dynamics in fractured reservoirs.

However, scalability remains a major limitation of PINNs due to the complexity of solving partial differential equations (PDEs) within a neural network framework. To address this, researchers should explore methods such as domain decomposition, transfer learning, and physics-guided pre-training, which could significantly improve computational efficiency and make PINNs applicable to large-scale reservoir simulations.

Surrogate models provide a practical alternative to high-cost, physics-based numerical simulations by using ML to approximate reservoir dynamics. Convolutional Neural Networks (CNNs) and intelligent agent-based models have demonstrated success in predicting well productivity, optimizing field development, and providing near real-time decision-making support.

However, further improvements are needed to enhance their adaptability and robustness across varying reservoir conditions. For example, integrating meta-learning techniques could enable surrogate models to quickly adapt to new reservoir conditions without requiring extensive retraining. Additionally, reinforcement learning-based optimization strategies could enhance their decision-making capabilities, particularly in production scheduling and enhanced oil recovery (EOR) strategies.

To address the identified challenges and enhance the effectiveness of oil and gas production prediction, future research should focus on:

(1) Future oil and gas production forecasting models must optimize efficiency and scalability to handle vast data and complex computations. Developing sparse neural networks and model quantization can minimize computational redundancy, making ML models efficient for edge devices in oilfields. Distributed computing and federated learning will enable collaborative model training across wells and fields, reducing data transmission costs while ensuring privacy. Knowledge distillation can transfer knowledge from large models to lightweight ones, enabling real-time forecasting on resource-limited devices.

(2) Integrating diverse data sources is crucial for enhancing prediction accuracy and model applicability. A comprehensive data platform should unify seismic, logging, production, and remote sensing data to create high-quality training datasets. Data augmentation and imputation algorithms, such as GANs and Bayesian interpolation, can correct missing and noisy data, improving completeness and generalization. Deep Fusion Neural Networks (DFNN) can integrate multimodal data (images, text, time series) to enhance predictive accuracy and robustness.

(3) Given the dynamic nature of oilfield production, edge computing deployment will enable local execution of ML models, minimizing cloud latency and enhancing real-time predictions. A self-adaptive learning framework will allow models to adjust dynamically based on real-time data streams, improving transient production forecasting. Efficient streaming data processing will enhance responsiveness to sensor data, increasing reliability and predictive precision.

(4) To reduce manual tuning and enhance adaptability, AutoML techniques, such as NAS and BO, should be applied for automatic model selection and hyperparameter optimization. Reinforcement learning-based model selection will allow adaptive learning strategies for varying reservoir conditions, further improving accuracy. Self-supervised learning can help models learn reservoir dynamics in an unsupervised setting, reducing reliance on extensive labeled data and improving adaptability.

(5) Physics-Informed Neural Networks (PINNs) offer high interpretability in reservoir modeling. Future research should optimize PINNs' computational efficiency and integrate transfer learning for rapid adaptation to different geological conditions. Multi-scale physical-data fusion can combine microscopic pore-scale flow models with macroscopic reservoir engineering models, making predictions more realistic. Generative AI models (e.g., Transformers) can simulate reservoir properties, improving predictions in data-scarce conditions and enhancing the adaptability of physics-driven ML models.

(6) To facilitate ML adoption in oilfield operations, SCADA-compatible ML platforms should integrate predictive models with production monitoring systems for automated forecasting and intelligent decision-making. Digital twin technology should be advanced to create real-time virtual oilfields for reservoir monitoring and optimization. Collaboration with oilfield service companies can establish AI-driven smart field optimization solutions, enabling direct ML deployment on-site, boosting efficiency, and promoting sustainable hydrocarbon development.

7. Conclusions

The integration of machine learning (ML) and artificial intelligence (AI) in oil and gas production forecasting has significantly enhanced prediction accuracy, efficiency, and interpretability. This study examined various approaches, including steady-state capacity forecasting, spatiotemporal sequence modeling, physics-informed neural networks (PINNs), and agent-based modeling, highlighting key advancements and challenges. The main conclusions are as follows:

(1) Traditional methods like linear regression (LR) and decline curve analysis (DCA) are limited in capturing nonlinear reservoir behavior, while ML models, such as Support Vector Machines (SVM), Random Forests (RF), and Gradient Boosting (GB), provide better accuracy. Feature selection and hyperparameter tuning techniques enhance model robustness, but challenges remain in handling dynamic production conditions.

(2) Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) excel in modeling production dynamics. Hybrid models (e.g., CNN-LSTM, BiGRU-AM, and Graph Convolutional Networks) improve prediction accuracy by integrating spatial and temporal learning. High computational costs hinder large-scale applications, necessitating more efficient architectures like sparse neural networks and model quantization.

(3) PINNs integrate physical laws into ML models, improving prediction reliability, particularly in noisy and sparse datasets. Scalability remains a challenge due to the complexity of solving partial differential equations (PDEs). Optimizations like domain decomposition and transfer learning can enhance efficiency.

(4) Agent-based models and surrogate ML models (e.g., CNN-based frameworks) offer efficient alternatives to high-cost numerical simulations. Meta-learning and reinforcement learning can improve model adaptability to new reservoir conditions.

(5) To further improve oil and gas production forecasting, future research should focus on enhancing computational efficiency and scalability by developing sparse neural networks, federated learning, and knowledge distillation for real-time applications. Multi-source data integration should be strengthened by combining seismic, logging, production, and remote sensing data, while utilizing GANs and Bayesian interpolation to handle missing and noisy data. Real-time prediction and adaptive learning can be improved through edge computing, streaming data processing, and self-adaptive learning frameworks to enhance responsiveness. Automated model optimization should incorporate AutoML, reinforcement learning, and self-supervised learning to reduce manual intervention and improve adaptability. Additionally, interpretability and physics constraints should be enhanced by improving PINNs through multi-scale physics-data fusion and generative AI for better modeling in data-scarce environments. Finally, industrial application and deployment should be advanced by integrating ML models with SCADA systems, developing digital twin technology, and fostering collaboration with oilfield service companies to enable real-world implementation and optimization of production forecasting models.

The convergence of ML, AI, and physics-based modeling is transforming production forecasting, improving efficiency and sustainability. Addressing challenges in computational cost, real-time adaptability, and deployment will drive further advancements. Future innovations in hybrid modeling, scalable architectures, and automated learning will enhance predictive accuracy and industrial applicability, ensuring cost-effective and optimized oilfield development.

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

Funding: The authors are thankful for financial support from the National Natural Science Foundation of China (Grant Nos. 52122402, 42090024 and U24B6001), Shandong Provincial Natural Science Foundation (Grant No. ZR2022JQ23). The authors would also like to appreciate the reviewers and editors whose critical suggestions were very helpful to this article.

Data Availability Statement: All data used in this study are obtained from previously published literature, which has been appropriately cited in the references section.

Acknowledgments: We would like to express our gratitude to the reviewers and editors for their critical suggestions, which were instrumental in improving this article. Additionally, we appreciate the support and resources provided by China University of Petroleum, the State Key Laboratory of Deep Oil and Gas, as well as the Institute of Geology.

Conflicts of Interest: Author Yuqing Yang was employed by the company Daqing Oilfield Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Sherwin, E.D.; Rutherford, J.S.; Zhang, Z.; Chen, Y.; Wetherley, E.B.; Yakovlev, P.V.; Berman, E.S.F.; Jones, B.B.; Cusworth, D.H.; Thorpe, A.K.; et al. US Oil and Gas System Emissions from Nearly One Million Aerial Site Measurements. *Nature* **2024**, *627*, 328–334. [CrossRef] [PubMed]
2. Zheng, J.; Wang, C.; Liang, Y.; Liao, Q.; Li, Z.; Wang, B. Deeppipe: A Deep-Learning Method for Anomaly Detection of Multi-Product Pipelines. *Energy* **2022**, *259*, 125025. [CrossRef]
3. Obite, C.P.; Chukwu, A.; Bartholomew, D.C.; Nwosu, U.I.; Esiaba, G.E. Classical and Machine Learning Modeling of Crude Oil Production in Nigeria: Identification of an Eminent Model for Application. *Energy Rep.* **2021**, *7*, 3497–3505. [CrossRef]

4. Yuan, S.; Han, H.; Wang, H.; Luo, J.; Wang, Q.; Lei, Z.; Xi, C.; Li, J. Research Progress and Potential of New Enhanced Oil Recovery Methods in Oilfield Development. *Pet. Explor. Dev.* **2024**, *51*, 963–980. [CrossRef]
5. Fan, J.; Qu, X.; Wang, C.; Lei, Q.; Cheng, L.; Yang, Z. Natural Fracture Distribution and a New Method Predicting Effective Fractures in Tight Oil Reservoirs in Ordos Basin, NW China. *Pet. Explor. Dev.* **2016**, *43*, 806–814. [CrossRef]
6. Wang, C.; Dong, T.; Chen, L.; Zhu, G.; Chen, Y. Multi-Objective Optimization Approach for Permanent Magnet Machine via Improved Soft Actor–Critic Based on Deep Reinforcement Learning. *Expert Syst. Appl.* **2025**, *264*, 125834. [CrossRef]
7. Wang, C.; Zheng, J.; Liang, Y.; Wang, B.; Klemeš, J.J.; Zhu, Z.; Liao, Q. Deeppipe: An Intelligent Monitoring Framework for Operating Condition of Multi-Product Pipelines. *Energy* **2022**, *261*, 125325. [CrossRef]
8. Bardy, G.; Biver, P.; Caumon, G.; Renard, P. Oil Production Uncertainty Assessment by Predicting Reservoir Production Curves and Confidence Intervals from Arbitrary Proxy Responses. *J. Pet. Sci. Eng.* **2019**, *176*, 116–125. [CrossRef]
9. Agboada, D.K.; Ahmadi, M. Production Decline and Numerical Simulation Model Analysis of the Eagle Ford Shale Oil Play. *OnePetro* **2013**, *221*, SPE-165315-MS. [CrossRef]
10. Aydin, H.; Boppano, N.; Yurukcu, M.; Liu, S.; Yegin, C.; Temizel, C. A Comprehensive Review of RTA/DCA Methods in Unconventional Reservoirs. In Proceedings of the 2022 SPE Western Regional Meeting, Bakersfield, CA, USA, 26–28 April 2022; SPE: Bakersfield, CA, USA, 2022; Volume 176, pp. 116–125.
11. Tang, H.; Zhang, B.; Liu, S.; Li, H.; Huo, D.; Wu, Y.-S. A Novel Decline Curve Regression Procedure for Analyzing Shale Gas Production. *J. Nat. Gas Sci. Eng.* **2021**, *88*, 103818. [CrossRef]
12. Chen, Z.; Yu, W.; Liang, J.-T.; Wang, S.; Liang, H. Application of Statistical Machine Learning Clustering Algorithms to Improve EUR Predictions Using Decline Curve Analysis in Shale-Gas Reservoirs. *J. Pet. Sci. Eng.* **2022**, *208*, 109216. [CrossRef]
13. Bao, A.; Gildin, E.; Huang, J.; Coutinho, E.J. Data-Driven End-To-End Production Prediction of Oil Reservoirs by EnKF-Enhanced Recurrent Neural Networks. *OnePetro* **2020**, *111*, SPE-199005-MS. [CrossRef]
14. Du, J.; Zheng, J.; Liang, Y.; Lu, X.; Klemeš, J.J.; Varbanov, P.S.; Shahzad, K.; Rashid, M.I.; Ali, A.M.; Liao, Q.; et al. A Hybrid Deep Learning Framework for Predicting Daily Natural Gas Consumption. *Energy* **2022**, *257*, 124689. [CrossRef]
15. Wang, C.; Zheng, J.; Du, J.; Wang, G.; Klemeš, J.J.; Wang, B.; Liao, Q.; Liang, Y. Weather Condition-Based Hybrid Models for Multiple Air Pollutants Forecasting and Minimisation. *J. Clean. Prod.* **2022**, *352*, 131610. [CrossRef]
16. Al Dwood, R.; Meng, Q.; Ibrahim, A.-W.; Yahya, W.A.; Alareqi, A.; AL-Khulaidi, G. A Novel Hybrid ANN-GB-LR Model for Predicting Oil and Gas Production Rate. *Flow Meas. Instrum.* **2024**, *100*, 102690. [CrossRef]
17. Duan, H.; Wang, G.; Song, Y.; Chen, H. A Novel Time-Delay Multivariable Grey Model and Its Application in Predicting Oil Production. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109505. [CrossRef]
18. Anderson, R.N. "Petroleum Analytics Learning Machine" for Optimizing the Internet of Things of Today's Digital Oil Field-to-Refinery Petroleum System. In Proceedings of the 2017 IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; Volume 3, pp. 4542–4545.
19. Bai, W.-P.; Cheng, S.-Q.; Guo, X.-Y.; Wang, Y.; Guo, Q.; Tan, C.-D. Oilfield Analogy and Productivity Prediction Based on Machine Learning: Field Cases in PL Oilfield, China. *Pet. Sci.* **2024**, *21*, 2554–2570. [CrossRef]
20. Awad, M.; Khanna, R. Support Vector Regression. In *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*; Apress: New York, NY, USA, 2015; Volume 5, pp. 67–80.
21. Lu, C.; Jiang, H.; Yang, J.; Wang, Z.; Zhang, M.; Li, J. Shale Oil Production Prediction and Fracturing Optimization Based on Machine Learning. *J. Pet. Sci. Eng.* **2022**, *217*, 110900. [CrossRef]
22. Bi, K.; Qiu, T. An Intelligent SVM Modeling Process for Crude Oil Properties Prediction Based on a Hybrid GA-PSO Method. *Chin. J. Chem. Eng.* **2019**, *27*, 1888–1894. [CrossRef]
23. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
24. Guo, X.-Q.; Wang, L.-S.; Rong, S.-X.; Guo, T.-M. Viscosity Model Based on Equations of State for Hydrocarbon Liquids and Gases. *Fluid Phase Equilibria* **1997**, *139*, 405–421. [CrossRef]
25. Golzari, A.; Haghighat Sefat, M.; Jamshidi, S. Development of an Adaptive Surrogate Model for Production Optimization. *J. Pet. Sci. Eng.* **2015**, *133*, 677–688. [CrossRef]
26. Yang, X.; Dindoruk, B.; Lu, L. A Comparative Analysis of Bubble Point Pressure Prediction Using Advanced Machine Learning Algorithms and Classical Correlations. *J. Pet. Sci. Eng.* **2020**, *185*, 106598. [CrossRef]
27. Sarker, I.H.; Kayes, A.S.M.; Watters, P. Effectiveness Analysis of Machine Learning Classification Models for Predicting Personalized Context-Aware Smartphone Usage. *J. Big Data* **2019**, *6*, 57. [CrossRef]
28. Morozov, A.D.; Popkov, D.O.; Duplyakov, V.M.; Mutalova, R.F.; Osiptsov, A.A.; Vainshtein, A.L.; Burnaev, E.V.; Shel, E.V.; Paderin, G.V. Data-Driven Model for Hydraulic Fracturing Design Optimization: Focus on Building Digital Database and Production Forecast. *J. Pet. Sci. Eng.* **2020**, *194*, 107504. [CrossRef]
29. Zhong, Z.; Sun, A.Y.; Wang, Y.; Ren, B. Predicting Field Production Rates for Waterflooding Using a Machine Learning-Based Proxy Model. *J. Pet. Sci. Eng.* **2020**, *194*, 107574. [CrossRef]

30. Zhao, H.; Hao, X. Location Decision of Electric Vehicle Charging Station Based on a Novel Grey Correlation Comprehensive Evaluation Multi-Criteria Decision Method. *Energy* **2024**, *299*, 131356. [CrossRef]
31. Hou, G.; Wang, J.; Fan, Y. Multistep Short-Term Wind Power Forecasting Model Based on Secondary Decomposition, the Kernel Principal Component Analysis, an Enhanced Arithmetic Optimization Algorithm, and Error Correction. *Energy* **2024**, *286*, 129640. [CrossRef]
32. Wang, K.; Qi, X.; Liu, H.; Song, J. Deep Belief Network Based k-Means Cluster Approach for Short-Term Wind Power Forecasting. *Energy* **2018**, *165*, 840–852. [CrossRef]
33. Mei, Z.; Shi, Z. On LASSO for High Dimensional Predictive Regression. *J. Econom.* **2024**, *242*, 105809. [CrossRef]
34. Kim, S.; Hong, Y.; Lim, J.-T.; Kim, K.H. Improved Prediction of Shale Gas Productivity in the Marcellus Shale Using Geostatistically Generated Well-Log Data and Ensemble Machine Learning. *Comput. Geosci.* **2023**, *181*, 105452. [CrossRef]
35. Gu, J.-W.; Zhou, X.; Wang, S.-L. A Prediction Method for Oilfield Production Based on Lasso Algorithm. *Sci. Technol. Eng.* **2020**, *20*, 10759–10763.
36. Lian, Z.-Q. Prediction of Oil Well Production Based on Deep Learning. Master's Thesis, Xinjiang University, Ürümqi, China, 2020.
37. Shevade, S.K.; Keerthi, S.S.; Bhattacharyya, C.; Murthy, K.K. Improvements to the SMO Algorithm for SVM Regression. *IEEE Trans. Neural Netw.* **2000**, *11*, 1188–1193. [CrossRef] [PubMed]
38. Han, S.; Wang, Z.; El-Meligy, M.; Alnowibet, K.A. Nonlinear Dynamic Analysis of the FG-TPMS Double-Curved Panels: Introducing SVM-DNN-RF Algorithm to Predict Nonlinear Dynamic Information. *Aerosp. Sci. Technol.* **2025**, *158*, 109785. [CrossRef]
39. Ibrahim, M.; Wattenbarger, R.A. Analysis of Rate Dependence in Transient Linear Flow in Tight Gas Wells. *OnePetro* **2006**, *31*, 1118–1163. [CrossRef]
40. Wevill, J.; Bromhead, A.; Evans, K.; Yarus, J.; John, C.M. Relative Performance of Support Vector Machine, Decision Trees, and Random Forest Classifiers for Predicting Production Success in US Unconventional Shale Plays. *Adv. Subsurf. Data Anal.* **2022**, *1*, 31–62. ISBN 978-0-12-822295-9.
41. He, Y.-W.; He, Z.-Y.; Tang, Y.; Qin, J.-Z.; Song, J.-J.; Wang, Y. Evaluation and Prediction of Shale Gas Well Production Based on Machine Learning. *Pet. Drill. Technol.* **2021**, *43*, 518–524. [CrossRef]
42. Liu, X.-P.; Deng, J.; Yang, P.-L. Fractured Horizontal Well Productivity Prediction in Tight Oil Reservoirs Based on K-Means and SVR. *Comput. Digit. Eng.* **2023**, *51*, 1949–1953.
43. Chi, R.-Q.; Zhang, W.-J.; Li, G.-R.; Wang, Y.-M. Method for Predicting Shale Oil Productivity in Chang 7 Formation of Ordos Basin Using Well-Logging Data. In Proceedings of the 2023 International Conference on Oil and Gas Exploration and Development, Wuhan, China, 19–21 September 2023; Volume 1, pp. 423–430.
44. Yang, Y.; Zhao, H.-X.; Yuan, S.-B. Oil and Gas Well Production Prediction Model Based on GWO-SVM Algorithm. *Energy Environ. Prot.* **2024**, *46*, 178–183. [CrossRef]
45. Al-Shami, T.M.; Jufar, S.R.; Kumar, S.; Abdulelah, H.; Abdullahi, M.B.; Al-Hajri, S.; Negash, B.M. A Comprehensive Review of Interwell Interference in Shale Reservoirs. *Earth-Sci. Rev.* **2023**, *237*, 104327. [CrossRef]
46. Shi, L. A New Proven Reserves Prediction Method Based on Random Forest Algorithm. *China Pet. Explor.* **2023**, *28*, 167–172.
47. Lei, M.; Qi, T.-J.; Yin, S.; Zhou, C.-Y.; Li, D.; Shu, X.-Y.; Luo, Y.; Shi, H.-Z. Optimization of Shale Gas Fracturing Parameters in Luzhou Block Based on Random Forest and Particle Swarm Optimization (RF-PSO). *Nat. Gas Technol. Econ.* **2023**, *17*, 53–61.
48. Yan, H.; He, Z.; Gao, C.; Xie, M.; Sheng, H.; Chen, H. Investment Estimation of Prefabricated Concrete Buildings Based on XGBoost Machine Learning Algorithm. *Adv. Eng. Inform.* **2022**, *54*, 101789. [CrossRef]
49. Bikmukhametov, T.; Jäschke, J. Oil Production Monitoring Using Gradient Boosting Machine Learning Algorithm. *IFAC-Pap.* **2019**, *52*, 514–519. [CrossRef]
50. Liao, L.-L.; Li, G.-S.; Zeng, Y.-J.; Song, X.-Z.; Gao, Q.-C.; Zhou, J. Prediction of Low-Permeability Reservoir Production Based on Data Mining and Machine Learning Techniques. *J. Yangtze Univ. (Nat. Sci. Ed.)* **2023**, *20*, 91–97. [CrossRef]
51. Liu, X.; Zhang, T.; Yang, H.; Qian, S.; Dong, Z.; Li, W.; Zou, L.; Liu, Z.; Wang, Z.; Zhang, T.; et al. Explainable Machine Learning-Based Method for Fracturing Prediction of Horizontal Shale Oil Wells. *Processes* **2023**, *11*, 2520. [CrossRef]
52. Hashemi Fath, A.; Madanifar, F.; Abbasi, M. Implementation of Multilayer Perceptron (MLP) and Radial Basis Function (RBF) Neural Networks to Predict Solution Gas-Oil Ratio of Crude Oil Systems. *Petroleum* **2020**, *6*, 80–91. [CrossRef]
53. Rahmanifard, H.; Gates, I.; Asl, A.S. Comparison of Machine Learning and Statistical Predictive Models for Production Time Series Forecasting in Tight Oil Reservoirs. *SPE* **2022**, *10*, 15. [CrossRef]
54. Wang, S.; Chen, Z.; Chen, S. Applicability of Deep Neural Networks on Production Forecasting in Bakken Shale Reservoirs. *J. Pet. Sci. Eng.* **2019**, *179*, 112–125. [CrossRef]
55. Liu, W.; Liu, W.; Gu, J.; Ji, C.; Sui, G. Study on Reservoir Interwell Connectivity Using Kalman Filter and Artificial Neural Network. *Oil Gas Geol. Recovery Effic.* **2020**, *27*, 118–124. [CrossRef]

56. Jiang, Y.; Zhang, H.; Zhang, K.; Wang, J.; Han, J.; Cui, S.; Zhang, L.; Zhao, H.; Liu, P.; Song, H. Waterflooding Interwell Connectivity Characterization and Productivity Forecast with Physical Knowledge Fusion and Model Structure Transfer. *Water* **2023**, *15*, 218. [CrossRef]
57. Sharma, A.; Suryawanshi, B.; Mohanty, B.; Sawarkar, A.N. Comparison of Artificial Neural Network and Response Surface Methodology for Evaluation of the Predictive Capability of Bio-Oil Yield from Pyrolysis of Mangifera Indica Wood Sawdust. *Fuel* **2023**, *338*, 127251. [CrossRef]
58. Wu, C.; Fu, X.; Xia, Y. Spare Part Demand Forecasting Using PSO Trained Quantile Regression Neural Network. *Comput. Ind. Eng.* **2025**, *200*, 110841. [CrossRef]
59. Göncü, A.; Kuzubaş, T.U.; Saltoğlu, B. Predicting Oil Prices: A Comparative Analysis of Machine Learning and Image Recognition Algorithms for Trend Prediction. *Financ. Res. Lett.* **2024**, *67*, 105874. [CrossRef]
60. Dong, Y.; Song, L.; Zhao, Q.; Ding, Z.; Qiu, L.; Lu, C.; Chen, G. A Physics-Guided eXtreme Gradient Boosting Model for Predicting the Initial Productivity of Oil Wells. *Geoenergy Sci. Eng.* **2023**, *231*, 212402. [CrossRef]
61. Daviran, M.; Maghsoudi, A.; Ghezelbash, R. Optimized AI-MPM: Application of PSO for Tuning the Hyperparameters of SVM and RF Algorithms. *Comput. Geosci.* **2025**, *195*, 105785. [CrossRef]
62. Dai, X.; Liu, L.; Cheng, Z. Elucidating Price Variability Drivers in Highway Electromechanical Equipment Using CV Predictions with PSO-XGBoost. *Alex. Eng. J.* **2024**, *109*, 754–767. [CrossRef]
63. Xue, Y.; Zhang, C. A Novel Importance-Guided Particle Swarm Optimization Based on MLP for Solving Large-Scale Feature Selection Problems. *Swarm Evol. Comput.* **2024**, *91*, 101760. [CrossRef]
64. Wang, K.; Li, H.; Wang, J.; Jiang, B.; Bu, C.; Zhang, Q.; Luo, W. Predicting Production and Estimated Ultimate Recoveries for Shale Gas Wells: A New Methodology Approach. *Appl. Energy* **2017**, *206*, 1416–1431. [CrossRef]
65. Liu, P.; Mu, Z.; Wang, W.; Liu, P.; Hao, M.; Liu, J. A New Combined Solution Model to Predict Water Cut in Water Flooding Hydrocarbon Reservoirs. *Int. J. Hydrogen Energy* **2017**, *42*, 18685–18690. [CrossRef]
66. Min, C.; Wen, G.; Gou, L.; Li, X.; Yang, Z. Interpretability and Causal Discovery of the Machine Learning Models to Predict the Production of CBM Wells after Hydraulic Fracturing. *Energy* **2023**, *285*, 129211. [CrossRef]
67. Du, J.; Zheng, J.; Liang, Y.; Ma, Y.; Wang, B.; Liao, Q.; Xu, N.; Ali, A.M.; Rashid, M.I.; Shahzad, K. A Deep Learning-Based Approach for Predicting Oil Production: A Case Study in the United States. *Energy* **2024**, *288*, 129688. [CrossRef]
68. Fang, W.; Chen, S.; Shi, S. Dynamic Characteristics and Real-Time Control of a Particle-to-sCO_2 Moving Bed Heat Exchanger Assisted by BP Neural Network. *Energy* **2022**, *256*, 124597. [CrossRef]
69. Zemlianova, K.; Bose, A.; Rinzel, J. Dynamical Mechanisms of How an RNN Keeps a Beat, Uncovered with a Low-Dimensional Reduced Model. *Sci. Rep.* **2024**, *14*, 26388. [CrossRef] [PubMed]
70. Luo, R.; Li, Y.; Guo, H.; Wang, Q.; Wang, X. Cross-Operating-Condition Fault Diagnosis of a Small Module Reactor Based on CNN-LSTM Transfer Learning with Limited Data. *Energy* **2024**, *313*, 133901. [CrossRef]
71. Tian, Y.; Ju, B. A Shale Gas Production Decline Prediction Model Based on BP Neural Network Improved by Genetic Algorithm. *China Sci. Pap.* **2016**, *11*, 1710–1715.
72. Zhou, J.; Zhang, H.; Wang, M. Application of Machine Learning Methods Constrained by Physical Models in Shale Oil Production Prediction. *Appl. Math. Mech.* **2021**, *42*, 881–890.
73. Liu, B.; Ma, Y.; Yasin, Q.; Wood, D.A.; Sun, M.; Gao, S.; Bai, Y. Characterization of Lacustrine Shale Oil Reservoirs Based on a Hybrid Deep Learning Model: A Data-Driven Approach to Predict Lithofacies, Vitrinite Reflectance, and TOC. *Mar. Pet. Geol.* **2025**, *174*, 107309. [CrossRef]
74. Hou, C. Production Prediction Method for New Oil Wells in Oilfields Based on Long-Short Term Memory Neural Network. *Oil Gas Geol. Recovery Effic.* **2019**, *26*, 105–110. [CrossRef]
75. Song, X.; Liu, Y.; Xue, L.; Wang, J.; Zhang, J.; Wang, J.; Jiang, L.; Cheng, Z. Time-Series Well Performance Prediction Based on Long Short-Term Memory (LSTM) Neural Network Model. *J. Pet. Sci. Eng.* **2020**, *186*, 106682. [CrossRef]
76. Liu, W.; Liu, W.D.; Gu, J. Forecasting Oil Production Using Ensemble Empirical Mode Decomposition Based Long Short-Term Memory Neural Network. *J. Pet. Sci. Eng.* **2020**, *189*, 107013. [CrossRef]
77. Cui, C.; Qian, Y.; Wu, Z.; Lu, S.; He, J. Forecasting of Oil Production Driven by Reservoir Spatial–Temporal Data Based on Normalized Mutual Information and Seq2Seq-LSTM. *Energy Explor. Exploit.* **2023**, *42*, 444–461. [CrossRef]
78. Fan, D.; Sun, H.; Yao, J.; Zhang, K.; Yan, X.; Sun, Z. Well Production Forecasting Based on ARIMA-LSTM Model Considering Manual Operations. *Energy* **2021**, *220*, 119708. [CrossRef]
79. Cho, K.; van Merrienboer, B.; Bahdanau, D.; Bengio, Y. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. *arXiv* **2014**, arXiv:1409.1259. [CrossRef]
80. Li, X.; Ma, X.; Xiao, F.; Xiao, C.; Wang, F.; Zhang, S. Time-Series Production Forecasting Method Based on the Integration of Bidirectional Gated Recurrent Unit (Bi-GRU) Network and Sparrow Search Algorithm (SSA). *J. Pet. Sci. Eng.* **2022**, *208*, 109309. [CrossRef]

81. Gao, Y.; Tang, L.-H.; Wang, Z.-P.; Xie, X.-Q.; Xu, H.-T. Production Prediction Method of Oilfields Based on Recurrent Neural Network and Data Differencing. *China Offshore Oil Gas* **2023**, *35*, 126–136.
82. Li, X.; Ma, X.; Xiao, F.; Xiao, C.; Wang, F.; Zhang, S. Multistep Ahead Multiphase Production Prediction of Fractured Wells Using Bidirectional Gated Recurrent Unit and Multitask Learning. *SPE J.* **2023**, *28*, 381–400. [CrossRef]
83. Haritha, D.; Satyavani, N.; Ramesh, A. Generation of Missing Well Log Data with Deep Learning: CNN-Bi-LSTM Approach. *J. Appl. Geophys.* **2025**, *233*, 105628. [CrossRef]
84. Tian, Q.; Luo, W.; Guo, L. Water Quality Prediction in the Yellow River Source Area Based on the DeepTCN-GRU Model. *J. Water Process Eng.* **2024**, *59*, 105052. [CrossRef]
85. Alla, S.; Adari, S.K. Temporal Convolutional Networks: A Unified Approach to Action. *arXiv* **2019**, *1*, 257–295. [CrossRef]
86. Li, D.; Wang, Z.; Zha, W.; Wang, J.; He, Y.; Huang, X.; Du, Y. Predicting Production-Rate Using Wellhead Pressure for Shale Gas Well Based on Temporal Convolutional Network. *J. Pet. Sci. Eng.* **2022**, *216*, 110644. [CrossRef]
87. Zhen, Y.; Fang, J.; Zhao, X.; Ge, J.; Xiao, Y. Temporal Convolution Network Based on Attention Mechanism for Well Production Prediction. *J. Pet. Sci. Eng.* **2022**, *218*, 111043. [CrossRef]
88. Zhang, L.; Dou, H.; Wang, T.; Wang, H.; Peng, Y.; Zhang, J.; Liu, Z.; Mi, L.; Jiang, L. Production Prediction Method of Single Well in Water Flooding Oilfield Based on Integrated Temporal Convolutional Network Model. *Pet. Explor. Dev.* **2022**, *49*, 1150–1160. [CrossRef]
89. Zha, W.; Liu, Y.; Wan, Y.; Luo, R.; Li, D.; Yang, S.; Xu, Y. Forecasting Monthly Gas Field Production Based on the CNN-LSTM Model. *Energy* **2022**, *260*, 124889. [CrossRef]
90. Song, F.; Ding, H.; Wang, Y.; Zhang, S.; Yu, J. A Well Production Prediction Method of Tight Reservoirs Based on a Hybrid Neural Network. *Energies* **2023**, *16*, 2904. [CrossRef]
91. Li, W.; Wang, L.; Dong, Z.; Wang, R.; Qu, B. Reservoir Production Prediction with Optimized Artificial Neural Network and Time Series Approaches. *J. Pet. Sci. Eng.* **2022**, *215*, 110586. [CrossRef]
92. Zhou, Q.; Lei, Z.; Chen, Z.; Wang, Y.; Liu, Y.; Xu, Z.; Liu, Y. Shale Oil Production Predication Based on an Empirical Model-Constrained CNN-LSTM. *Energy Geosci.* **2023**, *5*, 102. [CrossRef]
93. Chen, G.; Tian, H.; Xiao, T.; Xu, T.; Lei, H. Time Series Forecasting of Oil Production in Enhanced Oil Recovery System Based on a Novel CNN-GRU Neural Network. *Geoenergy Sci. Eng.* **2024**, *233*, 212528. [CrossRef]
94. Kim, J.; Yang, H.; Choe, J. Robust Optimization of the Locations and Types of Multiple Wells Using CNN Based Proxy Models. *J. Pet. Sci. Eng.* **2020**, *193*, 107424. [CrossRef]
95. Fargalla, M.A.M.; Yan, W.; Deng, J.; Wu, T.; Kiyingi, W.; Li, G.; Zhang, W. TimeNet: Time2Vec Attention-Based CNN-BiGRU Neural Network for Predicting Production in Shale and Sandstone Gas Reservoirs. *Energy* **2024**, *290*, 130184. [CrossRef]
96. Pan, S.; Yang, B.; Wang, S.; Guo, Z.; Wang, L.; Liu, J.; Wu, S. Oil Well Production Prediction Based on CNN-LSTM Model with Self-Attention Mechanism. *Energy* **2023**, *284*, 128701. [CrossRef]
97. Zhou, G.; Guo, Z.; Sun, S.; Jin, Q. A CNN-BiGRU-AM Neural Network for AI Applications in Shale Oil Production Prediction. *Appl. Energy* **2023**, *344*, 121249. [CrossRef]
98. Kilic, U.; Karadag, O.O.; Ozyer, G.T. AGMS-GCN: Attention-Guided Multi-Scale Graph Convolutional Networks for Skeleton-Based Action Recognition. *Knowl.-Based Syst.* **2025**, *311*, 113045. [CrossRef]
99. Du, L.; Liu, Y.; Xue, L.; You, G. A Deep Learning Framework Using Graph Convolutional Networks for Adaptive Correction of Interwell Connectivity and Gated Recurrent Unit for Performance Prediction. *SPE Reserv. Eval. Eng.* **2022**, *25*, 815–831. [CrossRef]
100. Liu, G.; Shan, Z.; Meng, Y.; Akbar, T.A.; Ye, S. RDPGNet: A Road Extraction Network with Dual-View Information Perception Based on GCN. *J. King Saud Univ.—Comput. Inf. Sci.* **2024**, *36*, 102009. [CrossRef]
101. Du, E.; Liu, Y.; Xue, L.; Zhu, X.; Song, L. A Data-Driven Model for Production Prediction of Strongly Heterogeneous Reservoir under Uncertainty. *Geoenergy Sci. Eng.* **2023**, *223*, 211542. [CrossRef]
102. Xu, Z.; Leung, J.Y. Shale Gas Production Forecasting with Well Interference Based on Spatial-Temporal Graph Convolutional Network. *OnePetro* **2023**, *1*, 2421–2439.
103. Zhao, Z.; Chen, Y.; Zhang, Y.; Mei, G.; Luo, J.; Yan, H.; Onibudo, O.O. A Deep Learning Model for Predicting the Production of Coalbed Methane Considering Time, Space, and Geological Features. *Comput. Geosci.* **2023**, *173*, 105312. [CrossRef]
104. Prantikos, K.; Chatzidakis, S.; Tsoukalas, L.H.; Heifetz, A. Physics-Informed Neural Network with Transfer Learning (TL-PINN) Based on Domain Similarity Measure for Prediction of Nuclear Reactor Transients. *Sci. Rep.* **2023**, *13*, 16840. [CrossRef] [PubMed]
105. Jahani-nasab, M.; Bijarchi, M.A. Enhancing Convergence Speed with Feature Enforcing Physics-Informed Neural Networks Using Boundary Conditions as Prior Knowledge. *Sci. Rep.* **2024**, *14*, 23836. [CrossRef] [PubMed]
106. Roy, S.; Sarkar, D.R.; Annavarapu, C.; Roy, P.; Lecampion, B.; Valiveti, D.M. Adaptive Interface-PINNs (AdaI-PINNs) for Transient Diffusion: Applications to Forward and Inverse Problems in Heterogeneous Media. *Finite Elem. Anal. Des.* **2025**, *244*, 104305. [CrossRef]
107. Yang, F.; Liu, H.; Li, X.-X.; Cao, J.-X. PINN Neural Network Method for Solving the Forward and Inverse Problem of Time-Fractional Telegraph Equation. *Results Eng.* **2025**, *25*, 103997. [CrossRef]

108. Hu, H.; Qi, L.; Chao, X. Physics-Informed Neural Networks (PINN) for Computational Solid Mechanics: Numerical Frameworks and Applications. *Thin-Walled Struct.* **2024**, *205*, 112495. [CrossRef]
109. Wang, H.; Wang, M.; Chen, S.; Hui, G.; Pang, Y. A Novel Governing Equation for Shale Gas Production Prediction via Physics-Informed Neural Networks. *Expert Syst. Appl.* **2024**, *248*, 123387. [CrossRef]
110. Qi, S.; Sarris, C.D. Physics-Informed Neural Networks for Multiphysics Simulations: Application to Coupled Electromagnetic-Thermal Modeling. In Proceedings of the 2023 IEEE/MTT-S International Microwave Symposium—IMS, San Diego, CA, USA, 11–16 June 2023; Volume 1, pp. 166–169.
111. Almajid, M.M.; Abu-Al-Saud, M.O. Prediction of Porous Media Fluid Flow Using Physics Informed Neural Networks. *J. Pet. Sci. Eng.* **2022**, *208*, 109205. [CrossRef]
112. Alenezi, F.; Mohaghegh, S. A Data-Driven Smart Proxy Model for a Comprehensive Reservoir Simulation. In Proceedings of the 2016 4th Saudi International Conference on Information Technology (Big Data Analysis) (KACSTIT), Riyadh, Saudi Arabia, 6–9 November 2016; Volume 1, pp. 1–6.
113. Kristensen, M.; Chugunov, N.; Gisolf, A.; Biagi, M.; Dubost, F. Real-Time Formation Evaluation and Contamination Prediction Through Inversion of Downhole Fluid-Sampling Measurements. *SPE Reserv. Eval. Eng.* **2018**, *22*, 531–547. [CrossRef]
114. Hegde, C.; Daigle, H.; Millwater, H.; Gray, K. Analysis of Rate of Penetration (ROP) Prediction in Drilling Using Physics-Based and Data-Driven Models. *J. Pet. Sci. Eng.* **2017**, *159*, 295–306. [CrossRef]
115. Dong, Y.; Qiu, L.; Lu, C.; Song, L.; Ding, Z.; Yu, Y.; Chen, G. A Data-Driven Model for Predicting Initial Productivity of Offshore Directional Well Based on the Physical Constrained eXtreme Gradient Boosting (XGBoost) Trees. *J. Pet. Sci. Eng.* **2022**, *211*, 110176. [CrossRef]
116. Rodriguez, A.X.; Salazar, D.A. Methodology for the Prediction of Fluid Production in the Waterflooding Process Based on Multivariate Long–Short Term Memory Neural Networks. *J. Pet. Sci. Eng.* **2022**, *208*, 109715. [CrossRef]
117. Angeles, R.; Torres-Verdín, C.; Hadibeik, A.; Sepehrnoori, K. Estimation of Capillary Pressure and Relative Permeability from Formation-Tester Measurements Using Design of Experiment and Data-Weighing Inversion: Synthetic and Field Examples. *J. Pet. Sci. Eng.* **2010**, *75*, 19–32. [CrossRef]
118. Khazaeni, Y.; Mohaghegh, S.D. Intelligent Production Modeling Using Full Field Pattern Recognition. *SPE Reserv. Eval. Eng.* **2011**, *14*, 735–749. [CrossRef]
119. Shahkarami, A.; Mohaghegh, S. Applications of Smart Proxies for Subsurface Modeling. *Pet. Explor. Dev.* **2020**, *47*, 400–412. [CrossRef]
120. Aboaba, A.; Martinez, Y.; Mohaghegh, S.; Shahnam, M.; Guenther, C.; Liu, Y. Smart Proxy Modeling Application of Artificial Intelligence & Machine Learning in Computational Fluid Dynamics. *Natl. Energy Technol. Lab. (NETL)* **2020**, *12*, 725–739.

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.