

Perspective

# Insights into the Application of Machine Learning in Reservoir Engineering: Current Developments and Future Trends

Hai Wang  and Shengnan Chen \* 

Department of Chemical and Petroleum Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada

\* Correspondence: snchen@ucalgary.ca

**Abstract:** In the past few decades, the machine learning (or data-driven) approach has been broadly adopted as an alternative to scientific discovery, resulting in many opportunities and challenges. In the oil and gas sector, subsurface reservoirs are heterogeneous porous media involving a large number of complex phenomena, making their characterization and dynamic prediction a real challenge. This study provides a comprehensive overview of recent research that has employed machine learning in three key areas: reservoir characterization, production forecasting, and well test interpretation. The results show that machine learning can automate and accelerate many reservoirs engineering tasks with acceptable level of accuracy, resulting in more efficient and cost-effective decisions. Although machine learning presents promising results at this stage, there are still several crucial challenges that need to be addressed, such as data quality and data scarcity, the lack of physics nature of machine learning algorithms, and joint modelling of multiple data sources/formats. The significance of this research is that it demonstrates the potential of machine learning to revolutionize the oil and gas sector by providing more accurate and efficient solutions for challenging problems.

**Keywords:** machine learning; oil and gas industry; production forecasting; well test analysis; reservoir characterization; challenges



**Citation:** Wang, H.; Chen, S. Insights into the Application of Machine Learning in Reservoir Engineering: Current Developments and Future Trends. *Energies* **2023**, *16*, 1392. <https://doi.org/10.3390/en16031392>

Academic Editor: Prabir Daripa

Received: 1 December 2022

Revised: 17 January 2023

Accepted: 25 January 2023

Published: 30 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Reservoir engineering is an interdisciplinary field that integrates mechanics, geology, physics, mathematics and computers as research tools to economically recover hydrocarbon resources in underground formations. Modelling and estimating such multiphysics and multiscale systems with conventional analytical or numerical simulation tools inevitably encounter serious challenges and introduce several sources of uncertainty [1]. In addition, investigating reservoir behaviors usually involves very intractable inverse problems, which typically have multi-solutions and require more complex theories and sophisticated algorithms. In recent decades, a great number of sensor-based tools are being adopted in the fields to automatically collect substantial amounts of data every day [2]. It is possible to fully discover the information underlying such valuable data to solve bottlenecks; this is where machine learning (ML) comes into play. ML enables the system to automatically learn and improve from prior information without being explicitly programmed [3], which is widely used to extract the highly nonlinear multi-factor interaction between numerous inputs and outputs (i.e., regression), pattern recognition, computer vision (CV), natural language processing (NLP) and discovering the governing partial differential equations, etc. In fact, the state-of-the-art ML algorithms even achieved human-level-surpassing performance in some specific tasks, such as AlphaGo in game [4] and residual networks (ResNets) in image recognition [5].

With the improvement of computing architecture, ML can also efficiently extract underlying information from the exploding real-world data collected in petroleum industry applications. As a powerful data-driven technique, ML is being widely accepted to assist

and improve our understanding of drilling [6], production and reservoir areas [7]. In particular, ML has been most widely used in reservoir engineering and has achieved excellent results, such as the prediction of permeability, porosity and tortuosity [8], prediction of shale gas production [9], reservoir characterization [10], 3D digital cores reconstruction [11], well test interpretation [12], rapid production optimization of shale gas [13,14], well-log processing [15] and history matching [16]. In this work, a systematic review of the ML methods applied to the different reservoir fields with an emphasis on production forecast, well test analysis and reservoir characterization is presented by reviewing the relevant literature published in recent years. In addition, despite the tremendous success of ML algorithms, achieved in a variety of reservoir applications that are computationally prohibitive or cannot be well modelled by physics principles, there are still many associated challenges and opportunities left for the future, which will also be outlined in this work.

## 2. Application Status of ML in Reservoir Engineering

Based on the available data, ML algorithms can be categorized into supervised, unsupervised, semi-supervised and reinforcement learning [17,18]. In reinforcement learning, algorithms learn by interacting with the environment and receiving feedback in the form of rewards or punishments. Unsupervised learning is applied to discover underlying patterns and obtain a more abstract representation of unlabeled data. Semi-supervised learning is employed if obtaining labeled data is expensive or time-consuming, but large amounts of unlabeled data are available. The algorithm utilizes the labeled data to extract the information and then applies this knowledge to label the unlabeled data. By far, supervised machine learning is more prevalent in a wide range of reservoir engineering applications, where trainable parameters are progressively updated for regression or classification purposes based on input–output pairs. ML can automate and accelerate many reservoirs engineering tasks, resulting in more efficient and cost-effective operations, as summarized in Table 1, and has been found to have a high level of accuracy and can be used to improve predictions in areas where traditional techniques have failed. This paper will mainly review the current application status of supervised ML in three aspects: production prediction, well test analysis and reservoir characterization.

**Table 1.** Commonly used ML model in reservoir engineering.

Model	Application	Reference	Advantage
LR	Cumulative production forecast	[19]	Easy to implement and interpret
FCNN	Well test interpretation	[20]	Powerful non-linear mapping capacity; flexibility in various tasks; good generalization
XGBoost	Cumulative production forecast	[21]	High accuracy; robust; missing values are acceptable
RF	Water invasion pattern identification; properties estimation	[12]	Parallelizable; robust
SVM	Surface oil rates prediction; permeability estimation	[22,23]	Robust to noise; Effective for small datasets
GPR	Early oil and gas production	[24]	Probabilistic approach; uncertainty estimation; Prior knowledge included

Table 1. Cont.

Model	Application	Reference	Advantage
ARIMA	Production dynamics	[25,26]	Interpretability; applicable to multiple temporal patterns;
LSTM	Production dynamics; reservoir models identification;	[27–30]	Capture long-term dependencies; powerful non-linear mapping capacity
CNN	Well test interpretation; production dynamics	[28,31–33]	Capture spatial dependencies; translation invariance

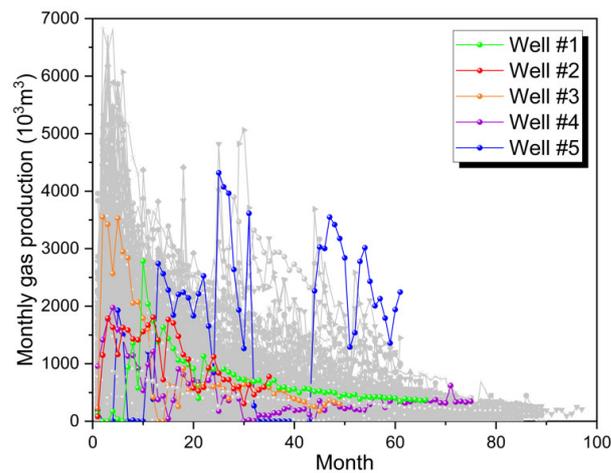
### 2.1. Production Prediction

Production forecasting plays an essential role in optimizing the construction strategy of wells, such as well drilling, stimulation and the enhanced hydrocarbon recovery processes [34].

A great number of investigations have employed ML to predict the cumulative production of oil and gas. Kong et al. [21] used the extreme gradient boost regressor (XGBoost) as the base model and then applied a linear regressor as the meta-model to ensemble multiple base models. A minor improvement ( $R^2$  improved from 0.79 to 0.8) was observed in predicting the cumulative production, but the bias of the stacked model was mitigated compared to that of the individual model. Wang et al. [19] trained and compared machine learning (ML) algorithms, including linear regression (LR), artificial neural networks (ANN), gradient-boosting decision trees (GBDT) and random forests (RF), to predict the first-year barrels of the oil equivalent (BOE) of infill wells. By employing the coefficient of determination ( $R^2$ ) and mean absolute error (MAE) as the metrics for model performance, ensemble methods (RF and GBDT) provide superior predictions than other methods, and the MAE and  $R^2$  were obtained as 0.79 and 31.8 kBbl, respectively. Support vector regression (SVR) and gaussian process regression (GPR) were used by Guo et al. [24] to predict early oil and gas production with a best  $R^2$  of 0.83, and sensitivity analysis indicated that fluid volume and total organic carbon (TOC) are the most essential features to predict the production of wells. With input variables GOR, upstream and downstream pressures, and choke size, support vector machine (SVM) and RF were implemented [22] to predict the surface oil rates in high gas oil ratio formations. All the data-driven models provided a better estimation ( $R^2 > 0.9$ ) than empirical correlation. Khan et al. [35] also employed artificial neuro fuzzy inference systems (ANFIS), SVM and ANN to investigate the oil production rate in artificial gas lift wells, and approximately 99% of variance in the data is explained.

In fact, predicting production dynamics with acceptable accuracy is of greater importance to reservoir operators, and production curve estimation is also a more challenging problem than predicting a single point (i.e., the cumulative production). As shown in Figure 1, the monthly well production in unconventional reservoirs fluctuates greatly due to the frequent shut-in operation, which grows the complexity of applying ML methods to deal with production sequences. The autoregressive strategy is widely used in time series analysis [36–38], i.e., to predict future values based on past values. For example, Duan et al. [25] integrated the autoregressive integral moving average (ARIMA) model and RTS smoothing to predict gas well production. Fan et al. [26] proposed a combined model to capture the linear component by ARIMA and the nonlinear part of the daily production sequence through a Long Short-Term Memory (LSTM) network separately. In addition, the specific architectures of recurrent neural networks (RNNs), such as LSTM and encoder–decoder architectures, render them inherently appropriate for modelling sequential information, such as monthly production and bottom hole pressures (BHP). A particle swarm optimization-assisted LSTM model was proposed by Song et al. [27] to infer the daily oil rate of fractured horizontal wells. Zha et al. [28] developed a CNN-LSTM

model to extract important features automatically and capture the sequence dependence, which was used to predict monthly natural gas production with an MAPE of 7.7%. Moreover, in another study [39], short-term forecasts of oil production were demonstrated by DeepAR and Prophet time series analysis. The results imply that ML approaches could fail in capturing long-term trends and the data-driven models should be retrained regularly. Zhong et al. [40] used a conditional deep convolutional generative neural network and the material balance method to develop a proxy model for the assessment of the production rate for waterflooding as a function of permeability field and time. The results indicate that the predicted values of the proposed model match excellently with the outcomes of the commercial simulator.



**Figure 1.** Monthly production of some wells in Duvernay Formation.

To summarize, ML has been widely adopted to estimate the early stage cumulative production and production dynamics of conventional/unconventional reservoirs, with  $R^2$  ranging from approximately 0.8 to 0.95, and neural networks are the most prevalent models due to their flexibility in processing different data formats and powerful non-linear mapping capacity. However, the production forecast problem involves data in multiple formats, such as time series and tabular data. Currently, processing data with multiple formats simultaneously can be challenging for ML models because different types of data have different characteristics. Therefore, integrating multiple data formats is an active area of research, and new architectures are being developed to address this limitation. In addition, the availability of real-world well data can be a major obstacle in developing accurate ML models for reservoir characterization because of the incomplete features collected and limited number of available samples, resulting in the poor generalization of the ML model.

## 2.2. Well Test Analysis

In well test analysis, an input impulse (typically a change in flow rate) is provided to the reservoir and then the corresponding response (typically a variation in pressure) is measured [41]. As is known to all, the response is governed by a set of underground properties such as porosity, permeability, well communication, formation damage (i.e., skin coefficient), boundary, and fracture geometries [42]. Based on the information collected and some analytical models that could link the response and petrophysical properties, it can be inferred that the model parameters/assumptions are analogous to subsurface conditions by applying optimal curve fitting [43,44]. ML also finds its feasibility in well test analysis. A study was conducted to implement CNN to estimate mobility ratios, dimensionless radius, well storage and skin effects in radial composite formations with input parameters of pressure buildup/drawdown data and pressure derivative information [31]. Leveraging the log–log plots (pressure and corresponding derivative), another researcher used CNN to model a dimensionless variable comprising permeability, storage and skin effect in infinite

reservoir [32]. Chu et al. [20] applied the multi-layer CNN and fully connected neural networks (FCNN) to classify well testing plots, and the mean  $F_1$  values were obtained as 0.91 and 0.81, respectively. Dong et al. [33] also used the one-dimensional CNN to interpret the well test data automatically. The well-trained model could be used to identify homogeneous, dual-porosity, radial composite and finite conductivity vertically fractured models and inverse the associated parameters with reasonable accuracy. Xue et al. [12] extracted the slopes of 40 segments to characterize the pressure derivative curve, and used these slopes to train RF to identify the water invasion pattern (bottom or edge water) in the gas reservoir. They also integrated the RF regressor and ensemble Kalman filter to investigate the permeability, aquifer size ratio and gas–water contact depth. Pandey et al. [45] used the genetic algorithm (GA) for feature selection and hyperparameter tuning to improve the performance of ANN in identifying and characterizing homogeneous reservoirs. The feasibility and superiority of GA-optimized ANN were demonstrated by applying the well-trained model to nine simulated cases and one real case. S. Wang and Chen [30] revealed that LSTM is competent in interpreting the correlation between pressure and flow rate data from hydraulic fractured tight reservoirs without the requirement of mathematical models. The authors concluded that the LSTM is able to capture well shut-in accurately and is more robust to noise. Nagaraj et al. [29] developed a Siamese neural network (SNN), which is composed of CNN and LSTM, for recognizing 14 different reservoir models. The SNN model achieved an accuracy of 93% in identifying the correct model as the top recommendation.

Well test analysis is heavily driven by both mathematics and data. It utilizes mathematical models based on physical principles to analyze data collected during well tests to estimate the performance of the well and the properties of the reservoir from which it is produced. ML has been increasingly applied in well test interpretation to improve the efficiency and accuracy, but the majority of current work has completely ignored the physical laws, leading to poor generalization and lack of interpretability.

### 2.3. Reservoir Characterization

Reservoir characterization is a heavily data-driven problem, which integrates seismic, logging and core analysis data to improve the understanding of subsurface properties such as porosity, saturation, permeability and pressure–volume–temperature [46]. The inherent heterogeneity of reservoirs makes the estimation of these subsurface properties very difficult [47], and relations between geophysical data and expected properties even fluctuate significantly from place to place. Therefore, conventional geostatistical approaches such as kriging and co-kriging could hardly identify the underlying correlation. Advances in ML algorithms facilitated the analysis of data from well logs, seismic surveys, and other sources to improve the understanding of the subsurface geology and fluid distribution.

A great number of studies have been carried out to infer reservoir properties inversely from seismic information because of the wide-range coverage. For instance, RNNs and Monte Carlo (MC) simulation were leveraged by Grana et al. [48] to classify facies from seismic data. The results reveal that RNNs perform better if the training set is sufficiently large, while MC can give the equivalent performance and quantify uncertainty better if the prior information is specified. Moreover, Liu et al. [49] also used RNNs architecture for seismic reservoir characterization. With input variables of seismic data, source wavelet and low-frequency prior porosity, the unsupervised convolutional neural network was developed by Feng et al. [50] to forecast porosity with limited labelled data. By introducing biased dropout and dropconnect strategies to address the overfitting problem, Liu et al. [10] extended the extreme learning machine (ELM) to simultaneously estimate several important properties such as lithofacies, shale content, porosity and saturation. The authors concluded that the proposed method has better generalization performance and a more efficient training process. Another work by Lee et al. [51] evaluated the statistical relationships between total organic carbon (TOC) in unconventional reservoirs and seismic parameters by combining ML and statistical rock physics. Chen et al. [52] proposed four physical

constraints (spatial, continuity, gradient and category) and incorporated them into the RF algorithm to predict reservoir quality. A significant improvement in prediction accuracy was observed in terms of the F1 score.

Although seismic data cover a large area and can provide properties assessments that span the entire reservoir, they are low-information-carrying [53–55]. Logging data have high resolution but low coverage, and core data provide the most accurate information and highest resolution, but their availability is limited. For instance, Katterbauer et al. [56] incorporated electrical and acoustic image logging data to classify fractures and estimate the fracture degree. With 100 shale scanning electron microscope images, Tian and Daigle [57] applied the automated object detection based on ML to identify and characterize microfractures. Several studies have focused on integrating various data sources to comprehensively characterize the reservoir, typically referred to as integration modelling [58,59]. Anifowose et al. [23] combined seismic data and wireline attributes to forecast the permeability by employing six state-of-the-art ML algorithms. The authors found that SVM outperformed the others, and the depth-matching strategy made a significant difference in estimating production capacities. Priezzhev and Stanislav [60] compared the generalization of conventional seismic inversion methods and several ML techniques using seismic attributes and well logs to predict reservoir properties. Investigating the distributions of lithological properties is useful to identify potential hydrocarbon-rich regions. Dixit et al. [61] utilized ascendant hierarchical clustering (AHC), self-organizing maps (SOM), ANN, and multi-resolution graph-based clustering (MRGC) with the best  $R^2$ s of 0.85, 0.74, 0.90 and 0.68, respectively, to predict lithofacies by integrating core data and well logs.

Overall, ML can be a powerful tool for reservoir characterization. However, seismic, logging and core analysis data are susceptible to errors or noises that may be introduced during the collection, aggregation or annotation stages, and small deviations in features could result in significant changes in ML estimates due to their inherent black-box nature. Therefore, high-quality data are required to make an accurate analysis and reliable decisions with the assistance of more advanced techniques.

### 3. Future Trends for ML in Reservoir Engineering

In spite of the remarkable achievement of ML in current reservoir applications, such as characterization and performance investigation, it is still in its infancy and has great potential for further improvement, as stated in Section 2.

#### 3.1. Data Quality and Quantity

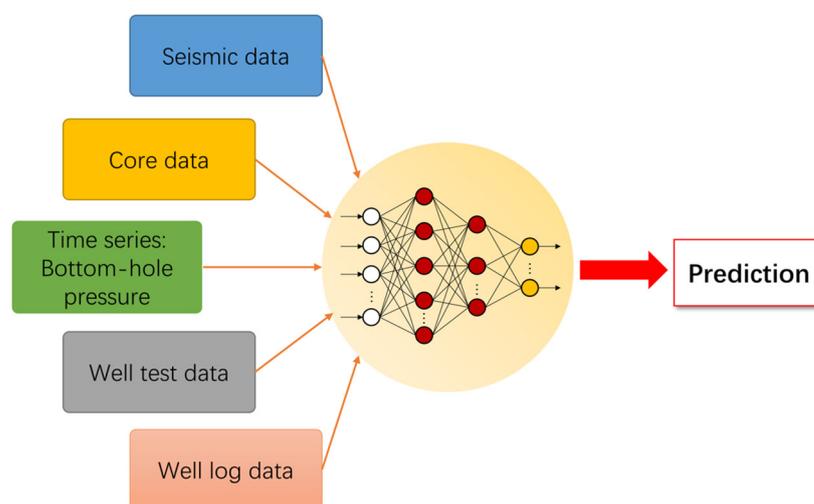
At present, most studies have been devoted to proposing new strategies to improve the performance of models, ignoring the fact that the quality of the model output depends not only on the model architectures, but also largely on the quality of the data [62]. Even most advanced ML algorithms cannot deliver reasonable results without guaranteed data quality and quantity. Data quality is not completely equal to data accuracy, but also comes with completeness and consistency [63]. Completeness denotes the absence of missing values and data consistency refers to whether the data records follow a uniform specification. The accuracy of data can be improved in the future with the assistance of more advanced sensors and mathematical algorithms (e.g., denoising and outlier detection). Completeness assurance requires the prior determination of what information must be collected in order to prevent the collection of too much redundant information, as well as the absence of many important parameters. Moreover, data consistency is most affected by human factors, and standardizing the data recording process, including the use of the same units and labeling methods, is crucial.

In addition, acquiring extremely high amounts of datasets is also very challenging in reservoir engineering due to the limitations of cost, technology and sharing policies. To overcome the shortage of samples, the construction of a data-sharing platform should be actively promoted. Moreover, the advent of few-shot learning, transfer learning and federated learning are likely to be promising solutions to partially address this challenge [64–66]. With

transfer learning, the model will be pre-trained on big data from a reservoir with similar conditions and then slightly tuned to enhance the model performance in the target reservoir. However, transfer learning may only be applicable on target domains with sufficient similarity, and some investigations have been conducted at the well scale [67]. Federated learning aims to train a model based on local training and parameter sharing without direct access to the data source, so data privacy and legal compliance are ensured. In addition, inspired by the capacity of humans to learn and generalize from small samples, few-shot learning aims to learn on scarce data with the assistance of strategies such as data augmentation, metric learning and meta-learning. In total, drawing reasonable inferences from sparse data is essential for reservoir systems, and current research is still in its early development, and more work is required in the future.

### 3.2. Fusion of Multiple Data Sources

In reservoir engineering, data at different scales, in different domains and different formats, are accumulated, e.g., there is a large number of computerized tomography (CT) scans at the core scale, time series data such as monthly fluid production, and a wide variety of tabular data at the well scale. Different data sources have different resolutions and can provide different information. For instance, seismic data are available regarding large areas and can provide an evaluation of petrophysical properties across the entire reservoir, but they only carry limited information. Logging data have relatively high resolution but suffer from sparse, site-specific data issues. Core data provide the most accurate information and highest resolution, but their acquisition is limited and often prohibitively expensive. Therefore, fusing multiple input information methods with ML models is expected to improve the generalization of the models, enhance the resolution of the reservoir modelling and alleviate the bottleneck of data scarcity. Therefore, reasonable stack strategies of various modules or multimodal learning architectures are anticipated to become available in the near future to process various inputs simultaneously. Figure 2 shows an illustration of the integration of multiple data sources in reservoir engineering.

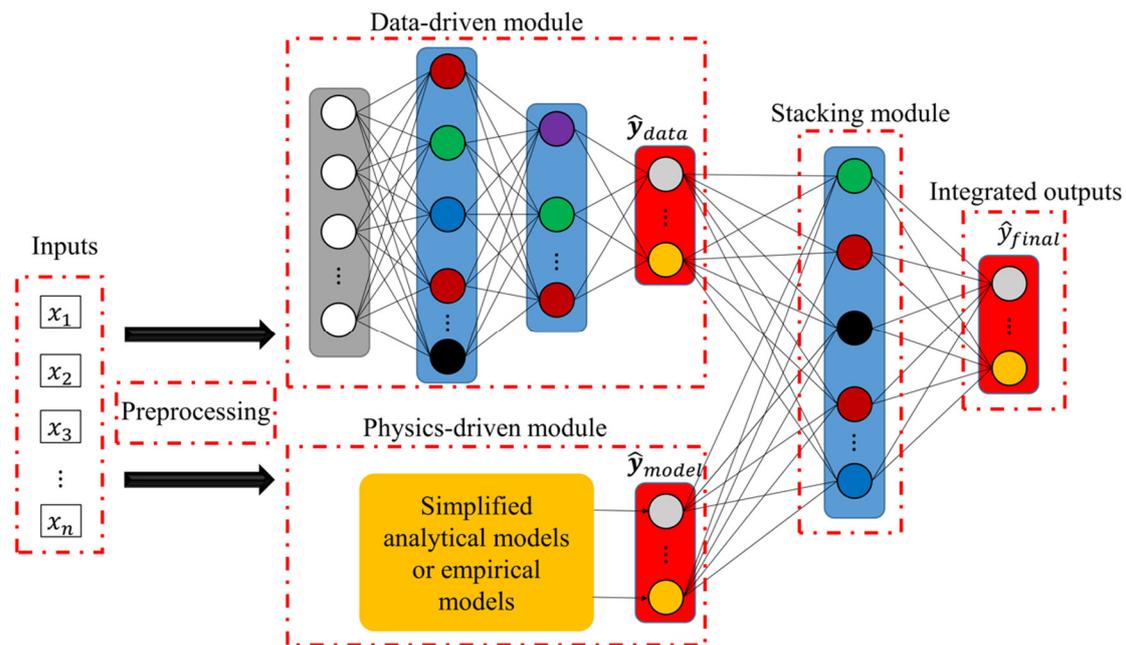


**Figure 2.** Illustration of integrating multiple data source.

### 3.3. Coupling Physics Laws with ML

Due to the capital-intensive nature of the oil and gas industry, a small deviation in decision making may result in a significant loss of manpower, resources and funds. Therefore, the interpretability of ML models is one of the most important and yet challenging issues, hindering the implementation of ML in the engineering field. The prior information of physical laws based on fundamental principles such as conservation laws, monotonicity and symmetry and empirical rules is completely discarded in ML practice in modelling physical phenomena, which results in poor generalization performance, especially in the

small data regime [68]. Therefore, even the state-of-the-art black-box methods are unable to provide physically consistent results and lack generalizability to out-of-bag samples [69]. Intuitively, applying extra physical constraints to a data-driven approach enables the ultimate model to benefit from both the data and the laws of physics and constraints the search space, thus granting more general inference with significantly fewer training samples than traditional methods, as well as ML. A simple coupling strategy is shown in Figure 3, using analytical/empirical models and machine learning algorithms (e.g., neural networks) to predict the desired variables, respectively, then stacking them and feeding them into a simple data-driven model to obtain the final output.



**Figure 3.** A simple strategy to integrate analytical/empirical models with data-driven models.

A more promising solution to effectively address this challenge could be the sophisticated physics-informed neural networks (PINN), in which automatic differentiation has been employed to calculate the derivatives of the neural network output with respect to input coordinates and model parameters. PINN has been successfully employed for data-driven partial differential equation solving, data-driven discovery governing equations and fitting a potential many-body energy surface [68,70,71], as such networks are constrained to respect any conservation principles, symmetries and differentiable property stemming from the physical laws [68]. Therefore, PINN is expected to have great potential for future applications in complex systems, such as reservoirs with large amounts of data and various applicable physical models.

#### 4. Conclusions

This study provides a comprehensive review of the ML approaches that have been employed in the field of reservoir engineering and highlights the application status and challenges. It is evident that ML has the potential to provide accurate results and improve reservoir engineering tasks, but there are still challenges to be addressed, and more research is needed to overcome these challenges. The key conclusions drawn from this study are listed below:

1. Machine learning (ML) techniques have numerous applications in reservoir engineering with acceptable accuracy, including the estimation of reservoir properties, well test interpretation, and investigation of production behaviors.

2. A variety of machine learning algorithms have been adopted in the field of reservoir engineering, among which neural networks (e.g., FCNN, RNN and its variant LSTM, CNNs) are the most popular models because of their powerful nonlinear mapping capacity and flexibility in addressing different data formats.
3. The current application of ML in reservoir engineering is still in its infancy, and further research is needed to enhance the ability to draw reliable inferences from sparse data and to develop strategies for integrating data from multiple sources/formats.
4. More attention should be given to the integration of physical laws with current data-driven models for the purpose of improving model interpretability and generalization, and PINN is a promising approach to address this problem.

**Author Contributions:** Conceptualization, H.W. and S.C.; writing—original draft preparation, H.W.; writing—review and editing, S.C. and H.W.; supervision, S.C.; project administration, S.C.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2020-05215) and Alliance Grant (ALLRP\548576-2019).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors gratefully acknowledge the financial support from the Discovery Grant (RGPIN-2020-05215) and the Alliance Grant (ALLRP\548576-2019) by Natural Sciences and Engineering Research Council of Canada (NSERC).

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

## References

1. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [[CrossRef](#)]
2. Choubey, S.; Karmakar, G.P. Artificial intelligence techniques and their application in oil and gas industry. *Artif. Intell. Rev.* **2021**, *54*, 3665–3683. [[CrossRef](#)]
3. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)] [[PubMed](#)]
4. Silver, D.; Schrittwieser, J.; Simonyan, K.; Antonoglou, I.; Huang, A.; Guez, A.; Hubert, T.; Baker, L.; Lai, M.; Bolton, A.; et al. Mastering the game of Go without human knowledge. *Nature* **2017**, *550*, 354–359. [[CrossRef](#)] [[PubMed](#)]
5. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
6. Noshi, C.I.; Schubert, J.J. The Role of Machine Learning in Drilling Operations; A Review. In Proceedings of the SPE/AAPG Eastern Regional Meeting, Pittsburgh, PA, USA, 7–11 October 2018. [[CrossRef](#)]
7. Purbey, R.; Parijat, H.; Agarwal, D.; Mitra, D.; Agarwal, R.; Pandey, R.K.; Dahiya, A.K. Machine learning and data mining assisted petroleum reservoir engineering: A comprehensive review. *Int. J. Oil Gas Coal Technol.* **2022**, *30*, 359–387. [[CrossRef](#)]
8. Graczyk, K.M.; Matyka, M. Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. *Sci. Rep.* **2020**, *10*, 21488. [[CrossRef](#)]
9. Hui, G.; Chen, S.; He, Y.; Wang, H.; Gu, F. Machine learning-based production forecast for shale gas in unconventional reservoirs via integration of geological and operational factors. *J. Nat. Gas Sci. Eng.* **2021**, *94*, 104045. [[CrossRef](#)]
10. Liu, X.; Ge, Q.; Chen, X.; Li, J.; Chen, Y. Extreme learning machine for multivariate reservoir characterization. *J. Pet. Sci. Eng.* **2021**, *205*, 108869. [[CrossRef](#)]
11. Bai, Y.; Berezovsky, V.; Popov, V. Digital core 3d reconstruction based on micro-CT images via a deep learning method. In Proceedings of the 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), Shenzhen, China, 23–25 May 2020; pp. 1–6.
12. Xue, L.; Gu, S.; Mi, L.; Zhao, L.; Liu, Y.; Liao, Q. An automated data-driven pressure transient analysis of water-drive gas reservoir through the coupled machine learning and ensemble Kalman filter method. *J. Pet. Sci. Eng.* **2022**, *208*, 109492. [[CrossRef](#)]
13. Wang, H.; Qiao, L.; Lu, S.; Chen, F.; Fang, Z.; He, X.; Zhang, J.; He, T. A Novel Shale Gas Production Prediction Model Based on Machine Learning and Its Application in Optimization of Multistage Fractured Horizontal Wells. *Front. Earth Sci.* **2021**, *9*, 726537. [[CrossRef](#)]
14. Qiao, L.; Wang, H.; Lu, S.; Liu, Y.; He, T. Novel Self-Adaptive Shale Gas Production Proxy Model and Its Practical Application. *ACS Omega* **2022**, *7*, 8294–8305. [[CrossRef](#)]
15. Wu, P.-Y.; Jain, V.; Kulkarni, M.S.; Abubakar, A. Machine learning-based method for automated well-log processing and interpretation. In Proceedings of the 2018 SEG International Exposition and Annual Meeting, Anaheim, CA, USA, 14–19 October 2018. [[CrossRef](#)]

16. Jo, H.; Pan, W.; Santos, J.E.; Jung, H.; Pycrz, M.J. Machine learning assisted history matching for a deepwater lobe system. *J. Pet. Sci. Eng.* **2021**, *207*, 109086. [[CrossRef](#)]
17. Ray, S. A quick review of machine learning algorithms. In Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 14–16 February 2019; pp. 35–39.
18. Mahesh, B. Machine learning algorithms—A review. *Int. J. Sci. Res.* **2020**, *9*, 381–386.
19. Wang, H.; Chen, Z.; Chen, S.; Hui, G.; Kong, B. Production forecast and optimization for parent-child well pattern in unconventional reservoirs. *J. Pet. Sci. Eng.* **2021**, *203*, 108899. [[CrossRef](#)]
20. Chu, H.; Liao, X.; Dong, P.; Chen, Z.; Zhao, X.; Zou, J. An Automatic Classification Method of Well Testing Plot Based on Convolutional Neural Network (CNN). *Energies* **2019**, *12*, 2846. [[CrossRef](#)]
21. Kong, B.; Chen, Z.; Chen, S.; Qin, T. Machine learning-assisted production data analysis in liquid-rich Duvernay Formation. *J. Pet. Sci. Eng.* **2021**, *200*, 108377. [[CrossRef](#)]
22. Ibrahim, A.F.; Al-Dhaif, R.; Elkatatny, S.; Al Shehri, D. Machine Learning Applications to Predict Surface Oil Rates for High Gas Oil Ratio Reservoirs. *J. Energy Resour. Technol.* **2022**, *144*, 1–19. [[CrossRef](#)]
23. Anifowose, F.; Abdulraheem, A.; Al-Shuhail, A. A parametric study of machine learning techniques in petroleum reservoir permeability prediction by integrating seismic attributes and wireline data. *J. Pet. Sci. Eng.* **2019**, *176*, 762–774. [[CrossRef](#)]
24. Guo, Z.; Wang, H.; Kong, X.; Shen, L.; Jia, Y. Machine Learning-Based Production Prediction Model and Its Application in Duvernay Formation. *Energies* **2021**, *14*, 5509. [[CrossRef](#)]
25. Duan, Y.; Wang, H.; Wei, M.; Tan, L.; Yue, T. Application of ARIMA-RTS optimal smoothing algorithm in gas well production prediction. *Petroleum* **2022**, *8*, 270–277. [[CrossRef](#)]
26. 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](#)]
27. 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](#)]
28. 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](#)]
29. Nagaraj, G.; Pillai, P.; Kulkarni, M. Deep Similarity Learning for Well Test Model Identification. In Proceedings of the SPE Middle East Oil & Gas Show and Conference, Event Canceled. 28 November–1 December 2021. [[CrossRef](#)]
30. Wang, S.; Chen, S. Application of the long short-term memory networks for well-testing data interpretation in tight reservoirs. *J. Pet. Sci. Eng.* **2019**, *183*, 106391. [[CrossRef](#)]
31. Li, D.; Liu, X.; Zha, W.; Yang, J.; Lu, D. Automatic well test interpretation based on convolutional neural network for a radial composite reservoir. *Pet. Explor. Dev.* **2020**, *47*, 623–631. [[CrossRef](#)]
32. Liu, X.; Li, D.; Yang, J.; Zha, W.; Zhou, Z.; Gao, L.; Han, J. Automatic well test interpretation based on convolutional neural network for infinite reservoir. *J. Pet. Sci. Eng.* **2020**, *195*, 107618. [[CrossRef](#)]
33. Dong, P.; Chen, Z.; Liao, X.; Yu, W. Application of deep learning on well-test interpretation for identifying pressure behavior and characterizing reservoirs. *J. Pet. Sci. Eng.* **2022**, *208*, 109264. [[CrossRef](#)]
34. Zhou, W.; Gupta, S.; Banerjee, R.; Poe, B.; Spath, J.; Thambaynagam, M. Production forecasting and analysis for unconventional resources. In Proceedings of the International Petroleum Technology Conference, Beijing, China, 26–28 March 2013.
35. Khan, M.R.; Alnuaim, S.; Tariq, Z.; Abdulraheem, A. Machine Learning Application for Oil Rate Prediction in Artificial Gas Lift Wells. In Proceedings of the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 18–21 March 2019. [[CrossRef](#)]
36. Wang, W.; Wong, A.K. Autoregressive Model-Based Gear Fault Diagnosis. *J. Vib. Acoust.* **2002**, *124*, 172–179. [[CrossRef](#)]
37. Lee, N.-U.; Shim, J.-S.; Ju, Y.-W.; Park, S.-C. Design and implementation of the SARIMA–SVM time series analysis algorithm for the improvement of atmospheric environment forecast accuracy. *Soft Comput.* **2018**, *22*, 4275–4281. [[CrossRef](#)]
38. Valipour, M.; Banihabib, M.E.; Behbahani, S.M.R. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J. Hydrol.* **2013**, *476*, 433–441. [[CrossRef](#)]
39. Tadjer, A.; Hong, A.; Bratvold, R.B. Machine learning based decline curve analysis for short-term oil production forecast. *Energy Explor. Exploit.* **2021**, *39*, 1747–1769. [[CrossRef](#)]
40. 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](#)]
41. Horne, R.N. *Modern Well Test Analysis*; Petroway Inc.: Palo Alto, CA, USA, 1995; p. 926.
42. Bourdet, D. *Well Test Analysis: The Use of Advanced Interpretation Models*; Elsevier: Amsterdam, The Netherlands, 2002.
43. Shun, L.; Feng-Rui, H.; Kai, Z.; Ze-Wei, T. Well Test Interpretation Model on Power-law Non-linear Percolation Pattern in Low-permeability Reservoirs. In Proceedings of the International Oil and Gas Conference and Exhibition in China, Beijing, China, 8–10 June 2010. [[CrossRef](#)]
44. Bourdet, D.; Ayoub, J.A.; Plard, Y.M. Use of Pressure Derivative in Well-Test Interpretation. *SPE Form. Eval.* **1989**, *4*, 293–302. [[CrossRef](#)]
45. Pandey, R.K.; Aggarwal, S.; Nath, G.; Kumar, A.; Vaferi, B. Metaheuristic algorithm integrated neural networks for well-test analyses of petroleum reservoirs. *Sci. Rep.* **2022**, *12*, 16551. [[CrossRef](#)]
46. Anifowose, F.A.; Labadin, J.; Abdulraheem, A. Hybrid intelligent systems in petroleum reservoir characterization and modeling: The journey so far and the challenges ahead. *J. Pet. Explor. Prod. Technol.* **2017**, *7*, 251–263. [[CrossRef](#)]

47. Vallabhaneni, S.; Saraf, R.; Priyadarshy, S. Machine-Learning-Based Petrophysical Property Modeling. In Proceedings of the SPE Europepec Featured at 81st EAGE Conference and Exhibition, London, UK, 3–6 June 2019.
48. Grana, D.; Azevedo, L.; Liu, M. A comparison of deep machine learning and Monte Carlo methods for facies classification from seismic data. *Geophysics* **2020**, *85*, WA41–WA52. [[CrossRef](#)]
49. Liu, M.; Nivlet, P.; Smith, R.; BenHasan, N.; Grana, D. Recurrent neural network for seismic reservoir characterization. In *Advances in Subsurface Data Analytics*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 95–116. [[CrossRef](#)]
50. Feng, R.; Hansen, T.M.; Grana, D.; Balling, N. An unsupervised deep-learning method for porosity estimation based on poststack seismic data Deep learning for porosity estimation. *Geophysics* **2020**, *85*, M97–M105. [[CrossRef](#)]
51. Lee, J.; Lumley, D.E.; Lim, U.Y. Improving total organic carbon estimation for unconventional shale reservoirs using Shapley value regression and deep machine learning methods. *AAPG Bull.* **2022**, *106*, 2297–2314. [[CrossRef](#)]
52. Chen, Y.; Zhao, L.; Pan, J.; Li, C.; Li, K.; Zhang, F.; Geng, J. Machine learning based deep carbonate reservoir characterization with physical constraints. In Proceedings of the 82nd EAGE Annual Conference & Exhibition, Online, 18–21 October 2021; pp. 1–5. [[CrossRef](#)]
53. Feng, R. Estimation of reservoir porosity based on seismic inversion results using deep learning methods. *J. Nat. Gas Sci. Eng.* **2020**, *77*, 103270. [[CrossRef](#)]
54. Saikia, P.; Baruah, R.D.; Singh, S.K.; Chaudhuri, P.K. Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models. *Comput. Geosci.* **2020**, *135*, 104357. [[CrossRef](#)]
55. Chaki, S.; Routray, A.; Mohanty, W.K. Well-Log and Seismic Data Integration for Reservoir Characterization: A Signal Processing and Machine-Learning Perspective. *IEEE Signal Process. Mag.* **2018**, *35*, 72–81. [[CrossRef](#)]
56. Katterbauer, K.; Qasim, A.; Al Shehri, A.; Al Zaidy, R. A Deep Learning Formation Image Log Classification Framework for Fracture Identification—A Study on Carbon Dioxide Injection Performance for the New Zealand Pohokura Field. In Proceedings of the SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 3–5 October 2022. [[CrossRef](#)]
57. Tian, X.; Daigle, H. Machine-learning-based object detection in images for reservoir characterization: A case study of fracture detection in shales. *Lead Edge* **2018**, *37*, 435–442. [[CrossRef](#)]
58. Le Ravalec, M.; Nouvelles, I.E.; Doligez, B.; Lerat, O. *Integrated Reservoir Characterization and Modeling*; IFP Energies Nouvelles: Rueil-Malmaison, France, 2014. [[CrossRef](#)]
59. Abdi, Y. Integrated reservoir characterization and modeling of one Iranian naturally fractured reservoir using laboratory and field data. In Proceedings of the SPE/EAGE Reservoir Characterization and Simulation Conference, Abu Dhabi, United Arab Emirates, 28–31 October 2007.
60. Priezzhev, I.; Stanislav, E. Application of Machine Learning Algorithms Using Seismic Data and Well Logs to Predict Reservoir Properties. In Proceedings of the 80th EAGE Conference and Exhibition 2018, Copenhagen, Denmark, 11–14 June 2018; pp. 1–5. [[CrossRef](#)]
61. Dixit, N.; McColgan, P.; Kusler, K. Machine Learning-Based Probabilistic Lithofacies Prediction from Conventional Well Logs: A Case from the Umiat Oil Field of Alaska. *Energies* **2020**, *13*, 4862. [[CrossRef](#)]
62. Polyzotis, N.; Zinkevich, M.; Roy, S.; Breck, E.; Whang, S. Data validation for machine learning. *Proc. Mach. Learn. Syst.* **2019**, *1*, 334–347.
63. Gudivada, V.; Apon, A.; Ding, J. Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. *Int. J. Adv. Softw.* **2017**, *10*, 1–20.
64. Konečný, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.T.; Bacon, D. Federated learning: Strategies for improving communication efficiency. *arXiv* **2016**, arXiv:1610.05492.
65. Weiss, K.; Khoshgoftaar, T.M.; Wang, D.D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [[CrossRef](#)]
66. Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a few examples: A survey on few-shot learning. *ACM Comput. Surv.* **2020**, *53*, 63. [[CrossRef](#)]
67. Razak, S.M.; Cornelio, J.; Cho, Y.; Liu, H.-H.; Vaidya, R.; Jafarpour, B. Transfer Learning with Recurrent Neural Networks for Long-Term Production Forecasting in Unconventional Reservoirs. *SPE J.* **2022**, *27*, 2425–2442. [[CrossRef](#)]
68. Raissi, M.; Perdikaris, P.; Karniadakis, G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]
69. Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; Kumar, V. Integrating physics-based modeling with machine learning: A survey. *arXiv* **2020**, arXiv:2003.04919.
70. Lu, L.; Meng, X.; Mao, Z.; Karniadakis, G.E. DeepXDE: A Deep Learning Library for Solving Differential Equations. *SIAM Rev.* **2021**, *63*, 208–228. [[CrossRef](#)]
71. Zhang, L.; Han, J.; Wang, H.; Saidi, W.; Car, R. End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018.

**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.