mathematics

MDPI

Article

Determination of Sequential Well Placements Using a Multi-Modal Convolutional Neural Network Integrated with Evolutionary Optimization

Seoyoon Kwon [1], Minsoo Ji [1,2], Min Kim [3], Juliana Y. Leung [4] and Baehyun Min [1,*]

[1] Department of Climate and Energy Systems Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea; sykwon024@ewhain.net (S.K.); msji048@ewhain.net (M.J.)
[2] POSCO International, 165, Convensia-daero, Yeonsu-gu, Incheon 21998, Republic of Korea
[3] Global E&P Technology Center, Korea National Oil Corporation, 305 Jongga-ro, Jung-gu, Ulsan 44538, Republic of Korea; m.kim@knoc.co.kr
[4] Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada; juliana2@ualberta.ca
* Correspondence: bhmin01@ewha.ac.kr

Abstract: In geoenergy science and engineering, well placement optimization is the process of determining optimal well locations and configurations to maximize economic value while considering geological, engineering, economic, and environmental constraints. This complex multi-million-dollar problem involves optimizing multiple parameters using computationally intensive reservoir simulations, often employing advanced algorithms such as optimization algorithms and machine/deep learning techniques to find near-optimal solutions efficiently while accounting for uncertainties and risks. This study proposes a hybrid workflow for determining the locations of production wells during primary oil recovery using a multi-modal convolutional neural network (M-CNN) integrated with an evolutionary optimization algorithm. The particle swarm optimization algorithm provides the M-CNN with full-physics reservoir simulation results as learning data correlating an arbitrary well location and its cumulative oil production. The M-CNN learns the correlation between near-wellbore spatial properties (e.g., porosity, permeability, pressure, and saturation) and cumulative oil production as inputs and output, respectively. The learned M-CNN predicts oil productivity at every candidate well location and selects qualified well placement scenarios. The prediction performance of the M-CNN for hydrocarbon-prolific regions is improved by adding qualified scenarios to the learning data and re-training the M-CNN. This iterative learning scheme enhances the suitability of the proxy for solving the problem of maximizing oil productivity. The validity of the proxy is tested with a benchmark model, UNISIM-I-D, in which four oil production wells are sequentially drilled. The M-CNN approach demonstrates remarkable consistency and alignment with full-physics reservoir simulation results. It achieves prediction accuracy within a 3% relative error margin, while significantly reducing computational costs to just 11.18% of those associated with full-physics reservoir simulations. Moreover, the M-CNN-optimized well placement strategy yields a substantial 47.40% improvement in field cumulative oil production compared to the original configuration. These findings underscore the M-CNN's effectiveness in sequential well placement optimization, striking an optimal balance between predictive accuracy and computational efficiency. The method's ability to dramatically reduce processing time while maintaining high accuracy makes it a valuable tool for enhancing oil field productivity and streamlining reservoir management decisions.

Keywords: well placement; multi-modal convolutional neural network; evolutionary optimization; UNISIM-I-D

check for updates

Academic Editors: Ioannis Tsoulos, Rui Wang and Shi Cheng

Received: 12 November 2024
Revised: 15 December 2024
Accepted: 23 December 2024
Published: 26 December 2024

Citation: Kwon, S.; Ji, M.; Kim, M.; Leung, J.Y.; Min, B. Determination of Sequential Well Placements Using a Multi-Modal Convolutional Neural Network Integrated with Evolutionary Optimization. *Mathematics* **2025**, *13*, 36. https://doi.org/10.3390/math13010036

Copyright: © 2024 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/).

MSC: 68T20

1. Introduction

In the field of geoenergy science and engineering (e.g., petroleum engineering, carbon capture and storage, and geothermal engineering), well placement optimization refers to the process of determining the optimal locations and configurations for wells in a reservoir to maximize economic value (e.g., cumulative oil production, cumulative CO_2 storage, cumulative heat energy extraction, and net present value) while considering various constraints [1–8]. These constraints include geological factors (e.g., reservoir properties, faults, and barriers), engineering limitations (e.g., well spacing, bottom-hole pressures, and flow rates), economic considerations (e.g., drilling costs, operational expenses, oil prices, and CO_2 prices), and environmental regulations. The process relies on reservoir simulation models that capture complex fluid flow involving thermal, hydrological, mechanical, and chemical interactions. This multi-million-dollar problem involves determining several key parameters: the number and type of wells (e.g., producers, injectors, and observation wells), well locations (e.g., x, y, and z coordinates), well types (e.g., vertical, horizontal, directional, and multilateral), well trajectories and completions, and operational parameters (e.g., flow rates, bottom-hole pressures, and injection/production schedules). The challenge lies in the complex, nonlinear relationships between variables, the computational intensity of reservoir simulations required for each evaluation, and the need to account for uncertainties in geological models and future economic scenarios. Therefore, the practical goal of well placement optimization is to develop efficient algorithms, often incorporating machine learning and artificial intelligence techniques, to find near-optimal solutions with minimal function evaluations while considering the inherent uncertainties and risks associated with subsurface operations.

Well placement optimization is essential to maximize the economic profits of oil and gas field development. Various optimization methods integrated with full-physics reservoir simulation (RS) have been used to determine well placement. For example, gradient-based algorithms, such as the adjoint-based method [1], simultaneous perturbation stochastic approximation [2], and stochastic simplest approximate gradient [9], have been integrated with RS to optimize various field development and operating plans. Non-gradient-based evolutionary optimization algorithms, such as particle swarm optimization (PSO) [3–5,10] and genetic algorithms (GA) [6–8], have also been widely employed because of their extensive population-based search capabilities. To find the global optimum, these population-based methods are considered powerful for generating high-quality solutions with stable convergence at affordable computational costs [11,12]. The geoenergy industry, particularly oil and gas, has widely adopted these evolutionary algorithms as primary optimization tools. Notably, these algorithms have been implemented in commercial software packages such as SLB's MEPO (2020.01) [13] and CMG's CMOST-AI (2023.12) [14]. However, their high computational cost owing to exhaustive simulation runs for objective function evaluations has hindered their implementation [15]. In addition, the stagnation of search performance in later generations during the evolutionary process requires improvement [16].

To alleviate the cost of an exhaustive search, various proxies have been employed to enhance the computational efficiency of these algorithms, while reducing the number of RS runs [17–19]. Machine learning techniques have been used to solve reservoir optimization problems, such as support vector machines [20,21] and Extreme Gradient Boosting [22,23]. These machine learning techniques can distinguish superior solutions from those that

violate constraints [15]. Recently, deep learning techniques such as artificial neural networks (ANNs) and convolutional neural networks (CNNs) have been demonstrated as partial or full substitutes for expensive reservoir simulators [24]. ANN has been actively adopted for well placement optimization [25–27]. However, ANNs cannot preserve the spatial features of large-scale problems because their input data are flattened into one-dimension [28]. Particularly, studies have demonstrated the excellent performance of CNNs in handling multi-dimensional data while preserving spatial features [28–30]. The multi-modal CNN (M-CNN), an enhanced CNN with multi-modal learning that functions by driving the fusion of input data from multiple sources [31], has been applied to predict cumulative oil production from near-wellbore static (e.g., permeability and porosity) and dynamic (e.g., pressure and saturation) properties [28,29].

Drawing inspiration from the M-CNN concept [29], this study seeks to enhance the solution quality of well placement problems by mitigating extrapolation effects and optimizing the learning data acquisition process. The performance of M-CNN is affected by the quantity and quality of learning data [32]. Considering the field development cost, the level of difficulty in applying the proxy has increased owing to the small ratio of the available data volume to the domain volume for well placement optimization problems [28]. To develop an efficient proxy, it is essential to balance accuracy and learning data acquisition costs. In the previous studies [28–30], generalization was achieved by randomly generating machine learning data, which increased the RS cost. For more efficiently managing the high computational cost of RSs associated with the acquisition of learning data, this study couples an M-CNN with an evolutionary algorithm (i.e., PSO). While PSO and GA are powerful for generating high-quality solutions, their high computational cost, resulting from exhaustive simulation runs, limits their practical application. In contrast, our approach integrates M-CNN with PSO and iterative learning, significantly reducing these costs while enhancing model accuracy. Another challenge in applying machine learning to well placement is the maximization problem that occurs when searching for the most productive region in a reservoir beyond the range of the learning data. Although features are captured well during machine learning interpolation, their predictability for extrapolation tends to deteriorate if the problem is nonlinear and complex [33–35]. We believe that the presence of PSO and iterative learning in the proposed framework enables the M-CNN to relieve the cost burden of RS. Constructing an M-CNN learned with PSO results and updating the proxy with iterative learning results are expected to allow the efficient selection of the optimal location with reasonable learning data volume for well placements. To the best of our knowledge, this is the first application of an evolutionary algorithm and iterative learning to obtain a dataset for a deep learning-based proxy model in the well placement optimization problem.

This study implements a hybrid workflow to sequentially determine locations that maximize hydrocarbon production. The proposed hybrid framework integrates M-CNN with PSO and iterative learning, addressing the high computational costs and accuracy limitations of traditional methods. This approach offers a more efficient and accurate solution for well placement optimization in dynamic, large-scale oil field developments. PSO was integrated to generate learning data within the framework of an M-CNN as a productivity estimator to determine the sequential placements of vertical oil production wells. Additionally, the concept of iterative learning was introduced to the M-CNN to mitigate the extrapolation problem. The M-CNN inputs spatial data (comprising static (e.g., permeability and porosity) and dynamic (e.g., pressure and oil saturation) reservoir properties) near a candidate well and outputs the oil productivity at that well location. Moreover, to consider the inter-well distance, auxiliary data (e.g., distances between the well and reservoir boundaries) are added as a 1D array to the full-connection stage of the

CNN. The M-CNN dataset was generated via a PSO-driven process by running the full-physics RSs of well-placing scenarios. The trained M-CNN is employed to evaluate the oil productivity at every candidate location. The results of the M-CNN are compared with the RS results for qualified (i.e., highly productive) well-placing scenarios to validate the model consistency. In addition, the M-CNN performance is compared with that of PSO to analyze the effect of using PSO to create datasets in the M-CNN. Expanding on the single vertical well placement optimization introduced by Kwon et al. [28], this study also enhances the methodology by enabling sequential well placement through a novel integration of the M-CNN with PSO and iterative learning techniques. This hybrid approach improves predictive accuracy and computational efficiency in dynamic reservoir scenarios, effectively addressing limitations identified in previous studies.

In brief, the key contributions of this study can be summarized as follows:

- Development of a Hybrid Optimization Framework:
  Introduced a novel hybrid framework integrating the M-CNN with PSO and iterative learning to optimize sequential well placement in reservoir management.
- Implementation of Iterative Learning for Enhanced Accuracy:
  Incorporated iterative learning to refine the training dataset, focusing on high-productivity scenarios to address extrapolation challenges and improve model generalization.
- Utilization of PSO for Efficient Data Generation:
  Leveraged PSO to generate learning data for the M-CNN, substantially reducing computational costs typically associated with full-physics reservoir simulations.
- Validation in Dynamic Reservoir Simulation Scenarios:
  Validated the proposed hybrid framework on the UNISIM-I-D benchmark model, demonstrating high predictive accuracy with reduced computational requirements.
- Pioneer Application of Integrated M-CNN Framework:
  To our knowledge, this study presents the first application of a hybrid M-CNN framework combined with PSO and iterative learning for dataset generation in sequential well placement optimization.

This paper is organized as follows: the literature review and scope of this study are introduced in Section 1; the methodology, including the theoretical background, structure of the M-CNN, and overall M-CNN workflow for sequential well placement optimization, is described in Section 2; the testing of the proposed methodology on a benchmark model UNISIM-I-D is presented in Section 3; and the conclusions based on the findings presented in Section 3 are summarized in Section 4.

2. Methodology

2.1. Well Placement Problem

Let $\mathbf{u} \in \mathbf{U}$ be a set of well locations in a given reservoir:

$$\mathbf{u} = [\mathrm{u}_1,\ \mathrm{u}_2,\ \cdots,\ \mathrm{u}_{N_w}] \tag{1}$$

where N_w is the number of wells, and $\mathrm{u}_i = (\mathrm{X}_i, \mathrm{Y}_i)$ is the coordinate in the Cartesian coordinate system. In this study, all wells in the reservoir were vertical and fully perforated. As a sequential well placement, u_i was drilled earlier than u_j where $1 \le i \le j \le N_w$.

$Q_o(\mathbf{u})$ is the expected field cumulative oil production at $\mathbf{u}$:

$$Q_o(\mathbf{u}) = \int_0^T q_o^t(\mathbf{u})dt, \tag{2}$$

where T is the total production time, and q_o^t is the oil production rate at time t.

The well placement optimization problem aims to determine the optimal set of well locations $\mathbf{u}^*$ that can yield the highest oil productivity.

$$\mathbf{u}^* = \underset{\mathbf{u}\in\mathbf{U}}{\operatorname{argmax}}(Q_o(\mathbf{u})). \tag{3}$$

To determine $\mathbf{u}^*$ via a cost-efficient approach, this study implemented a hybrid of an evolutionary optimization algorithm and deep learning algorithm. Figure 1 shows the workflow of the hybrid proxy that consists of PSO (Module 1), M-CNN (Module 2), and iterative learning (Module 3) for solving a sequential well placement problem. To verify the proxy performance, the full-physics RS results (e.g., oil production rate, water cut, and bottomhle pressure) are used as references (i.e., answers) throughout this study.

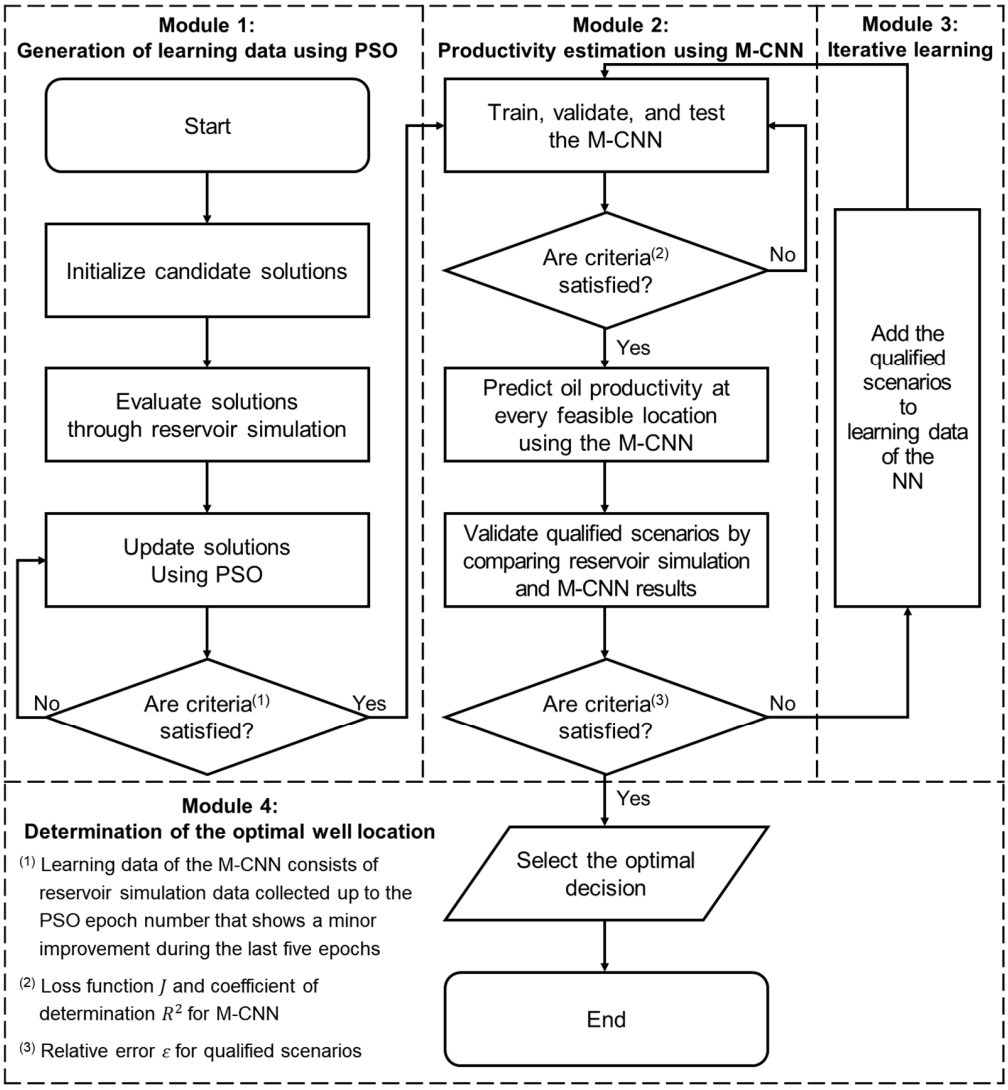


Figure 1. The flowchart of the proposed hybrid proxy to solve a sequential well placement problem. The optimum well locations are determined using the multi-modal convolutional neural network (M-CNN) incorporated with particle swarm optimization (PSO) and iterative learning.

2.2. Particle Swarm Optimization

PSO is a non-gradient population-based evolutionary optimization algorithm developed by Kennedy and Eberhart [10]. Figure 2a shows a schematic of the PSO. In PSO, the potential solutions are called particles, and a set of particles is referred to as a swarm [10]. Therefore, a swarm comprises the particles in the search space. The overall quality of

the particles, as well as the swarm, evolves during the evolutionary process (i.e., epochs). In this study, PSO searches for the optimal well location that reveals the highest field cumulative oil production by running a full-physics RS at feasible well locations. That is, the swarm evolves in the direction of increasing $Q_o(\mathbf{u})$.

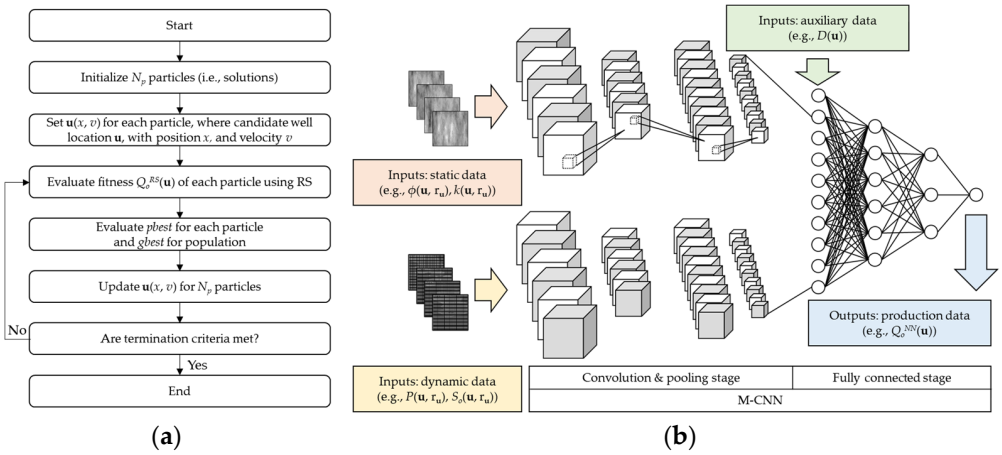


Figure 2. Schematic diagram of the PSO and M-CNN. (**a**) The PSO evolves solutions that are utilized as learning data of the M-CNN; (**b**) The M-CNN learns a pattern between inputs (i.e., petrophysical properties) and outputs (i.e., field cumulative oil production) extracted from the PSO solutions (modified from Chu et al. [29]).

Each particle corresponds to a candidate well location $\mathbf{u}$. A particle has velocity vector v and position vector x. The particle updates its v and x values based on the best individual and population solutions during each epoch.

v and x are updated as follows:

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 \left(x_{i,pbest}^k - x_i^k \right) + c_2 r_2 \left(x_{i,gbest}^k - x_i^k \right), \tag{4}$$

$$x_i^{k+1} = x_i^k + v_i^{k+1}, \tag{5}$$

where v_i^k and x_i^k are the velocity and position vector of the i-th $(1 \le i \le N_p)$ particle in the k-th $(1 \le k \le N_e)$ epoch, respectively. N_p and N_e are the population size (i.e., swarm size) and maximum number of epochs, respectively. ω, c_1, and c_2 are the inertia, cognitive, and social weight components, respectively. r_1 and r_2 are uniformly distributed random numbers sampled from [0.0, 1.0]. *pbest* is the historical best solution of the particle and *gbest* is the population historical best solution. Typically, ω is set in [0.4, 0.9]; c_1 and c_2 are in [1.0, 2.0] with reference to [10]. With reference to [36], the weight components are set as $\omega = 0.730$ and $c_1 = c_2 = 1.496$ throughout this study. PSO solves this well placement problem as an unconstrained problem in terms of the minimum inter-well distance because the M-CNN needs to learn that a short inter-well distance would yield poor oil production. During the optimization, the M-CNN learned well; thus, most solutions have sufficient inter-well distances. The use of PSO in this study is rooted in the desire to ensure the stability and reliability of the hybrid workflow. As this is the first attempt to integrate PSO with M-CNN and iterative learning, we chose to start with a well-established and proven optimization method. This approach minimizes the risks associated with using more complex or newer algorithms, providing a solid foundation for the workflow's performance and validation.

The proposed proxy employs PSO for two purposes. The first purpose is to generate learning data for the M-CNN. Specifically, the RS results obtained using PSO are used as learning data for the M-CNN: $(\mathbf{u}, Q_o^{RS}(\mathbf{u}))$, where the superscript RS refers to the full-physics reservoir simulation. This ensures the collection of learning data from more

hydrocarbon-prolific regions during the evolutionary process. Once the evolutionary process is complete, the best solution in the last epoch is considered as the optimal solution $\mathbf{u}^*$. Here, two stopping criteria for the evolutionary process are employed: N_e and the last epoch N_{es} with early stopping (ES) activation ($N_{es} \leq N_e$). The evolutionary process terminates early in the N_{es}-th epoch if the swarm quality improvement stagnates for the last five epochs. This early stopping criterion helps to reduce data acquisition costs, considering the high computational cost involved in field-scale RS. The second purpose of PSO in this study is to serve as a competitor for the developed proxy. In brief, the proxy results are compared with the PSO results with ES paying $N_p \times N_{es}$ RS runs, and with the simultaneous optimization of the field cumulative oil production $Q_o(\mathbf{u})$ and total computational cost.

2.3. Multi-Modal Convolutional Neural Network with Iterative Learning

Figure 2b illustrates the structure of the M-CNN developed for the sequential determination of well placements. This M-CNN inputs near-wellbore spatial properties (two static and two dynamic) and outputs the field cumulative oil production $Q_o^{NN}(\mathbf{u})$ at an arbitrary well location $\mathbf{u}$, where the superscript *NN* refers to a neural network (e.g., M-CNN in this study). Here, the spatial properties are porosity $\phi(\mathbf{u}, r_\mathbf{u})$, permeability $k(\mathbf{u}, r_\mathbf{u})$, pressure $P(\mathbf{u}, r_\mathbf{u})$, and oil saturation $S_o(\mathbf{u}, r_\mathbf{u})$, where $r_\mathbf{u}$ is the effective range from well $\mathbf{u}$ to an arbitrary reservoir location for determining the size of the 3D array data. Each location is assumed to be fully perforated. As sequential well placement enables the extraction of dynamic properties at each installation time, this study adopts multi-modal learning to input those properties. The features from the four spatial properties are extracted and upscaled during the convolutional and pooling stages. After convolution and pooling, the features are flattened and imported into the full-connection stage. In addition, the distance data $D(\mathbf{u})$ are also imported into the full-connection stage because they are scalar. $D(\mathbf{u})$ comprises the distance between the new well and reservoir boundaries and the distance between the new and existing wells.

The PSO results ($\mathbf{u}, Q_o^{RS}(\mathbf{u})$) are utilized as learning data for M-CNN. These learning data are split into training, validation, and test datasets for the efficient design and construction of the M-CNN. For each dataset, the performance of the M-CNN is assessed in terms of two indicators: loss function J and coefficient of determination R^2. Favored learning results are indicated by J values approaching zero from the right and R^2 values nearing one, signifying better model performance and fit.

The loss function J is designed to optimize the neural network's performance by balancing predictive accuracy with model generalizability, incorporating a regularization term that penalizes large weights to mitigate overfitting [32].

$$J = \mathrm{RMSE} + L(\mathrm{W}) = \sqrt{\frac{1}{N_d}\sum_{i=1}^{N_d}\left(Q_o^{RS}(\mathbf{u_i}) - Q_o^{NN}(\mathbf{u_i})\right)^2} + \frac{\alpha}{2}\|\mathrm{W}\|^2, \tag{6}$$

where RMSE is the root mean squared error, $L(\mathrm{W})$ is the L2 regulation term for weight array W, N_d is the number of data points, and α is the regularization coefficient. This study used a regularization coefficient of α = 0.01, based on our previous work [28] where this value effectively balanced under- and over-regularization.

The predictive accuracy of the M-CNN is quantified using R^2, where values close to 1 indicate strong agreement between predicted values Q_o^{NN} and reference values Q_o^{RS}. In other words, the high R^2 values demonstrate the M-CNN's capability to accurately predict cumulative oil production at well locations. These results highlight the effectiveness of iterative learning, and the quality of the dataset generated by PSO, which successfully captured representative scenarios for productive well placements.

$$R^2 = \frac{\left\{N_d\left(\sum_{i=1}^{N_d} Q_o^{RS}(\mathbf{u_i})Q_o^{NN}(\mathbf{u_i})\right) - \left(\sum_{i=1}^{N_d} Q_o^{RS}(\mathbf{u_i})\right)\left(\sum_{i=1}^{N_d} Q_o^{NN}(\mathbf{u_i})\right)\right\}^2}{\left\{N_d\sum_{i=1}^{N_d}\left(Q_o^{RS}(\mathbf{u_i})\right)^2 - \left(\sum_{i=1}^{N_d} Q_o^{RS}(\mathbf{u_i})\right)^2\right\}\left\{N_d\sum_{i=1}^{N_d}\left(Q_o^{NN}(\mathbf{u_i})\right)^2 - \left(\sum_{i=1}^{N_d} Q_o^{NN}(\mathbf{u_i})\right)^2\right\}}, \quad (7)$$

Also, the threshold for R^2, denoted as θ_{R^2}, is employed to judge the acceptance of the learning performance of the M-CNN. The learning results are accepted if all R^2 for training, validation, and test datasets are greater than θ_{R^2}.

The cumulative field oil production at every feasible well location (i.e., the active grid block in the reservoir model) $\mathbf{u} \in \mathbf{U}$ was rapidly estimated using the trained M-CNN. This rapid approximation of oil estimates for the entire reservoir domain generates a hydrocarbon productivity map similar to a quality map [2,26,37]. In the descending order of $Q_o^{NN}(\mathbf{u})$, the most highly productive N_{qual} grid blocks are nominated as qualified locations. To re-validate the robustness of the M-CNN, full-physics RS is performed for every qualified location to assess the similarity between the simulation and proxy results in terms of the relative error ε:

$$\varepsilon(\mathbf{u}) = \frac{\left|Q_o^{RS}(\mathbf{u}) - Q_o^{NN}(\mathbf{u})\right|}{Q_o^{RS}(\mathbf{u})} \times 100(\%), \quad (8)$$

$$\varepsilon_{avg} = \frac{1}{N_{qual}}\sum_{i=1}^{N_{qual}} \varepsilon(\mathbf{u}_i), \quad (9)$$

The grid block with the highest Q_o^{RS} is selected as the optimal well location, $\mathbf{u}^*$, if the average relative error, ε_{avg}, is acceptable. If ε_{avg} > threshold θ_ε, $(\mathbf{u}, Q_o^{RS}(\mathbf{u}))$ results of N_{qual} scenarios are added to the learning data, which are further split into training, validation, and test datasets, as shown in the flow chart (Figure 1). Subsequently, the M-CNN re-learns and yields an updated hydrocarbon productivity map. This iterative learning is performed until $\varepsilon_{avg} \leq \theta_\varepsilon$ is met.

Our hybrid approach synergistically combines the strengths of PSO and the M-CNN to address the limitations of standalone methods. PSO efficiently generates diverse, high-quality data, while M-CNN excels at preserving spatial features in high-dimensional datasets. The iterative learning process progressively refines the neural network model, enhancing its accuracy and robustness. This combination yields a more efficient and precise method for well placement optimization, overcoming the constraints of traditional techniques. By integrating these complementary approaches, our hybrid workflow mitigates common issues such as extrapolation errors and overfitting, which often plague standalone machine learning models or static optimization methods. Continuous dataset refinement through iterative learning ensures the model remains adaptive and increasingly accurate, leading to superior computational efficiency and predictive accuracy compared to conventional methods.

3. Results and Discussion

3.1. Field Description

An iterative learning-based proxy was applied to determine the locations of four vertical oil production wells for the UNISIM-I-D oil reservoir model at a primary production stage [38]. The original UNISIM-I-D is a benchmark reservoir model that encompasses the structural, facies, and petrophysical characteristics of the Namorado Field in Campos Basin, Brazil [39–41]. Table 1 summarizes the rock and fluid properties of UNISIM-I-D. Further details of the model properties are available in the original study [38]. This benchmark reservoir model was initially introduced to create a field development plan for

waterflooding, using 11 water injectors and 14 oil producers [42]. In this plan, waterflooding follows a four-year primary oil recovery (i.e., natural pressure drawdown) from four vertical oil wells installed at a four-month interval in a year, as designed in the original model [38]. Our analysis revealed that some of the primary production wells of the original model were installed in less prolific regions with low permeability, thus inspiring our motivation for well placement optimization for primary production. The optimization of the waterflooding plan will be investigated in future studies.

Table 1. Summary of the reservoir properties of the UNISIM-I-D used in this study.

Property	Value
Number of total grid blocks (X × Y × Z)	93,960 = 81 × 58 × 20
Number of active grid blocks	36,403
Grid block size (X × Y × Z), m × m × m	100 × 100 × 8
Depth of top layer, m	3000
Initial reservoir pressure, kg/cm^2	327.0
Porosity ($\mu \pm \sigma$), fraction	0.13 ± 0.08
Permeability ($\mu \pm \sigma$), mD	132.53 ± 184.95
Oil saturation ($\mu \pm \sigma$), fraction	0.56 ± 0.38

Figures 3–7 visualize 3D (Figure 3) and 2D distribution maps (Figures 4–7) of the top, middle, and bottom layers of the reservoir model for porosity, permeability, pressure, and oil saturation. The numbers of grid blocks in the X-, Y-, and Z-directions are 81, 58, and 20, respectively, for a total of 93,960 grid blocks in the UNISIM-I-D model. The number of active grid blocks is 36,403, which is the sum of the 2282, 2331, 2296, 802, 20, 2276, 2259, 2258, 2084, 2312, 2285, 2310, 1945, 16, 2308, 2200, 2221, 2250, 1883, and 65 active grid blocks from the top (1st) to the bottom (20th) layers. The inactive grid blocks are colored white.

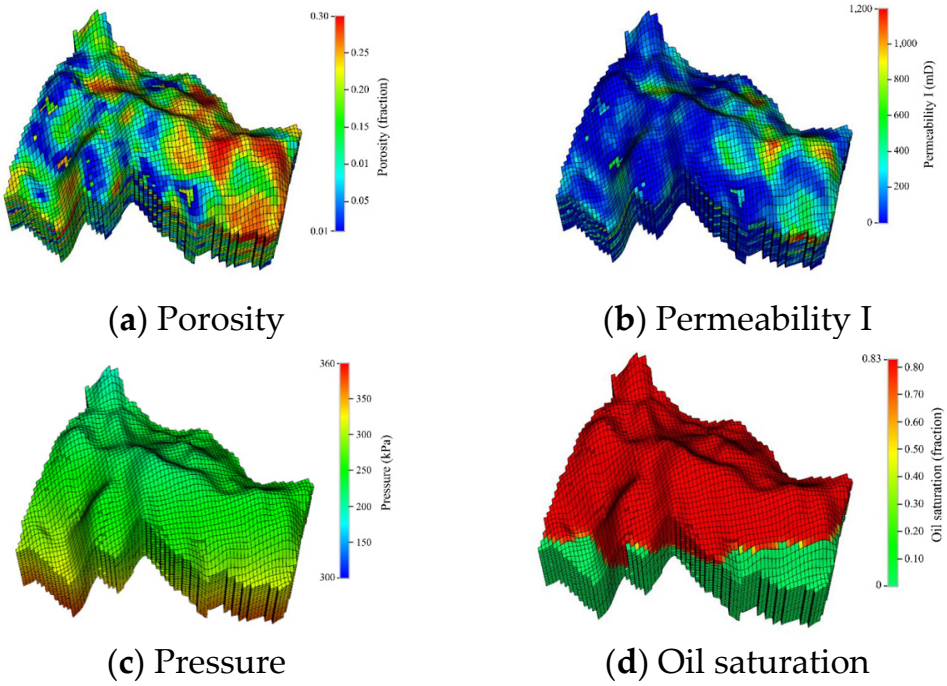
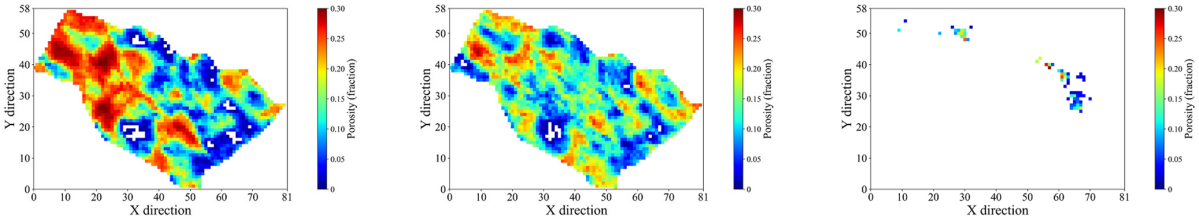
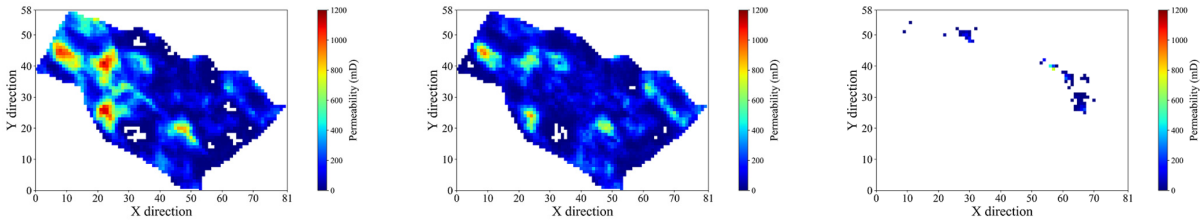


Figure 3. Three-dimensional distributions of petrophysical properties of the UNISIM-I-D. (**a**) Porosity; (**b**) permeability I; (**c**) pressure; (**d**) oil saturation. The total number of grid blocks is 93,960, with 36,403 of those being active.



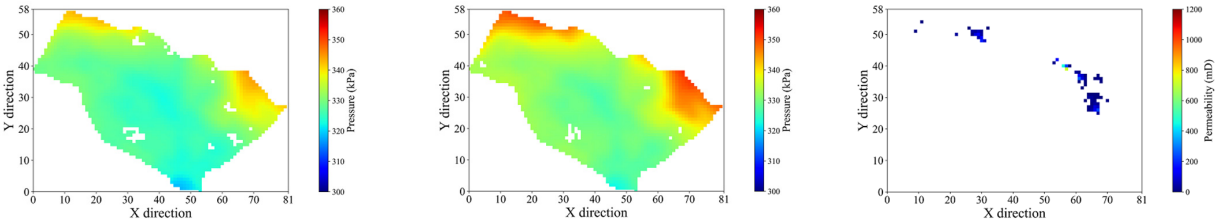
(**a**) 1st layer (**b**) 10th layer (**c**) 20th layer

Figure 4. Two-dimensional porosity distribution maps of the UNISIM-I-D. (**a**) The 1st layer (2282 active grid blocks); (**b**) 10th layer (2312 active grid blocks); (**c**) 20th layer (65 active grid blocks).



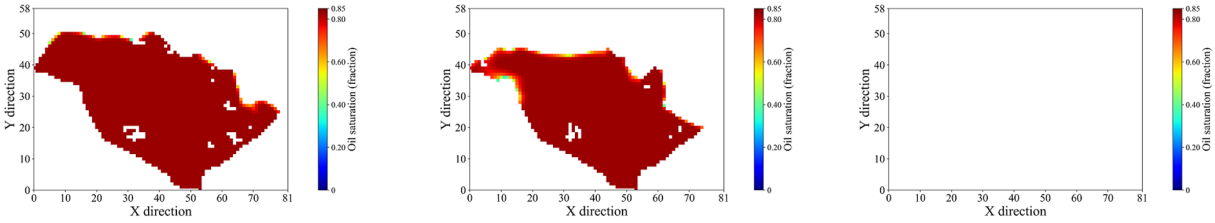
(**a**) 1st layer (**b**) 10th layer (**c**) 20th layer

Figure 5. Two-dimensional permeability distribution maps of the UNISIM-I-D. (**a**) The 1st layer (2282 active grid blocks); (**b**) 10th layer (2312 active grid blocks); (**c**) 20th layer (65 active grid blocks).



(**a**) 1st layer (**b**) 10th layer (**c**) 20th layer

Figure 6. Two-dimensional pressure distribution maps of the UNISIM-I-D. (**a**) The 1st layer (2282 active grid blocks); (**b**) 10th layer (2312 active grid blocks); (**c**) 20th layer (65 active grid blocks).



(**a**) 1st layer (**b**) 10th layer (**c**) 20th layer

Figure 7. Two-dimensional oil saturation distribution maps of the UNISIM-I-D. (**a**) The 1st layer (2282 active grid blocks); (**b**) 10th layer (2312 active grid blocks); (**c**) 20th layer (65 active grid blocks).

To maximize the four-year ($T = 4$) primary recovery, we initialized the well setting: all wells were removed from the reservoir model. This study focused on exploitation strategy during the four-year primary recovery period. Thereafter, this case study focuses on determining the locations of four primary production wells installed sequentially at the same four-month interval. Each vertical oil production well was fully perforated, except for the 4th, 5th, 14th, and 20th layers, where most of the grid blocks were inactive. The number of candidate well placement scenarios is equal to the number of active grid blocks in the top layer of the model. The numbers of candidates from the first to fourth production well placement problems are 2282, 2281, 2280, and 2279, respectively. In addition, the operating condition of each production well is a maximum surface liquid rate of 2000 m^3/day with a constraint of a minimum bottom-hole pressure of 150 kg/cm^2.

3.2. Generation of Learning Data Using Particle Swarm Optimization

Table 2 lists the experimental settings of PSO used for training the M-CNN (N_p = 20 and N_e = 20). The PSO algorithm parameters listed in Table 2 were adopted based on the settings referenced from the studies by Kennedy and Eberhart [10] and Clerc and Kennedy [36]. In this study, the PSO was performed using CMG's CMOST-AI (2023.12) [14]. Notably, the M-CNN learned the solutions (i.e., reservoir simulation results) accrued during the last five epochs, which exhibited a minor improvement in swarm quality. The early stopped epochs, N_{es}, were the 14th, 15th, 14th, and 6th epochs from the first to the fourth well placement problems, respectively. For example, Figure 8 shows box plots depicting the evolution of well placement scenarios for the first well placement. The overall quality of the learning data for the M-CNN improved progressively from the first generation to the last generation in the PSO process. Throughout the epochs, the highest value within the swarm increased, while its dispersion decreased. However, evolution stagnation was observed; the (near-) optima derived using PSO were identified in the 10th epoch, after which the evolution plateaued. Consequently, a total of 280 reservoir simulation results were accumulated up to the 14th epoch. Another limitation of the PSO is its redundancy. Up to the 14th epoch, 104 of the 280 results were duplicate well placement scenarios, indicating that the number of actual learning data obtained using the PSO (N_d) was 176 (=280 − 104). This redundancy occurred because the well coordinates were integer-type variables. Implementing an optimization algorithm to address redundancy in the solution set is expected to streamline learning data preparation for a machine learning technique.

Table 2. Experimental setting of the particle swarm optimization (PSO) algorithm used for training the M-CNN.

Parameter	Value
Population size, N_p	20
Maximum number of epochs, N_e	20
Inertia weight component, ω	0.730
Cognitive weight component, c_1	1.496
Social weight component, c_2	1.496

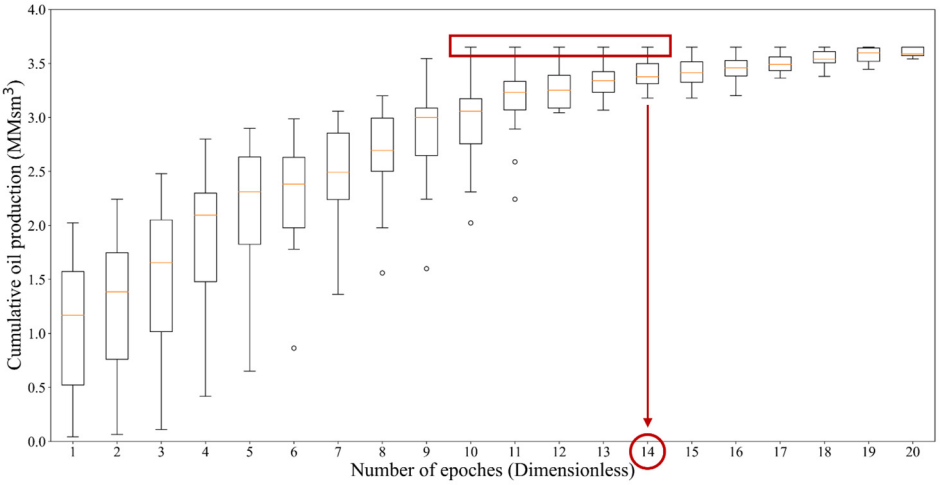


Figure 8. The box plots to show the evolution of well placement scenarios for the 1st well through PSO. Note that this study acquired the learning data for M-CNN through PSO. The PSO process initially improved solution quality, with increasing highest values and decreasing dispersion, but stagnated after the 10th epoch when (near-)optimal solutions were identified, concluding with 280 reservoir simulation results after 14 epochs. Small circles indicate outliers of the box plots.

3.3. Design of Multi-Modal Convolutional Neural Network

Table 3 summarizes the structure of the M-CNN used in the case study. As the initial reservoir model had homogeneous pressure and saturation, the first well placement case adopted a dual-modal M-CNN that inputs $\phi(\mathbf{u}, r_{\mathbf{u}})$ and $k(\mathbf{u}, r_{\mathbf{u}})$. For subsequent well placement cases, a quadmodal CNN with $\phi(\mathbf{u}, r_{\mathbf{u}})$, $k(\mathbf{u}, r_{\mathbf{u}})$, $P(\mathbf{u}, r_{\mathbf{u}})$, and $S_o(\mathbf{u}, r_{\mathbf{u}})$ was adopted. Here, $r_{\mathbf{u}} = 12$; thus, the 3D array size for each spatial property data was $25 \times 25 \times 20$.

Table 3. Structure of the M-CNN.

Parameter	Value
Number of input types	4
Type of inputs	Porosity ϕ, permeability k, pressure P, oil saturation S_o
Number of convolutional layers	2
Order of convolutional stage	Conv3D, BatchNormalization, AveragePooling3D
Number of hidden layers	3
Number of outputs	1
Type of outputs	Field cumulative oil production Q_o
Optimizer	Adam
Activation function	ReLU
Ratio of learning data splitting	Training 60%, validation 20%, test 20%

Each M-CNN comprises convolutional and fully connected stages. The convolutional stage consisted of two convolutional layers, two batch normalization layers, and two activation layers with ReLU [43]. Every element in each spatial input data array was normalized to [0.0,1.0]. Hyperparameters (e.g., the number of nodes in layers, convolutional filter, size of each filter, and type of pooling) were tuned to enhance the neural network performance. The learning stopping criteria during training and validation were early stopping and the maximum number of epochs. The patience in early stopping and maximum epoch number were 20 and 300, respectively. A threshold, $\theta_{R^2} = 0.75$, was used to evaluate the training, validation, and test performances, ensuring sufficient predictive accuracy before proceeding to subsequent steps. For iterative learning, $N_{qual} = 10$ was adopted from Kwon et al. [28], where it was a sufficient number to select representative scenarios. Additionally, θ_ε of 5% was set as the relative error tolerance considering Chu et al. [29], where their CNN model yielded an average relative error of 4.27%, with the higher threshold accounting for variability and ensuring robustness.

In this study, every M-CNN was built using the Keras deep learning library and its back-end engine TensorFlow in Python 3.11 environments, as in our previous studies [28,29]. All computations were executed using the following computer hardware specifications: Intel (R) Core (TM) i7–10700 CPU @ 2.9 GHz with 64.0 GB RAM and NVIDIA GeForce RTX 3060 Ti 8 GB.

3.4. Determination of the Well Locations and Performan ce Evaluation

For each sequential well placement, its learning data were divided into training (60%), validation (20%), and test (20%) sets, a commonly used ratio in machine learning [44]. A similar ratio was also employed in the previous studies [28–30], where it proved to be both reliable and effective in training proxy models for well placement optimization. Then, oil productivity at all feasible locations was evaluated using the learned proxy. Thereafter, the top ten qualified scenarios were selected in descending order of field cumulative oil production estimated using the proxy Q_o^{NN}. Table 4 summarizes the performance of the proxy for the learning data and the ε_{avg} values of the top ten qualified scenarios. All R^2 values exceed an θ_{R^2} of 0.75, confirming satisfactory learning performance. In addition, three cases (excluding the fourth well) were sufficiently trained using iterative learning. The relatively low R^2 value for the 1st well was attributed to the limited dataset available for training during its evaluation. As the iterative process progresses and more wells

were included, the dataset expanded, enabling the model to better generalize and reduce prediction uncertainty. This results in higher R^2 values for subsequent wells, reflecting improved predictive accuracy.

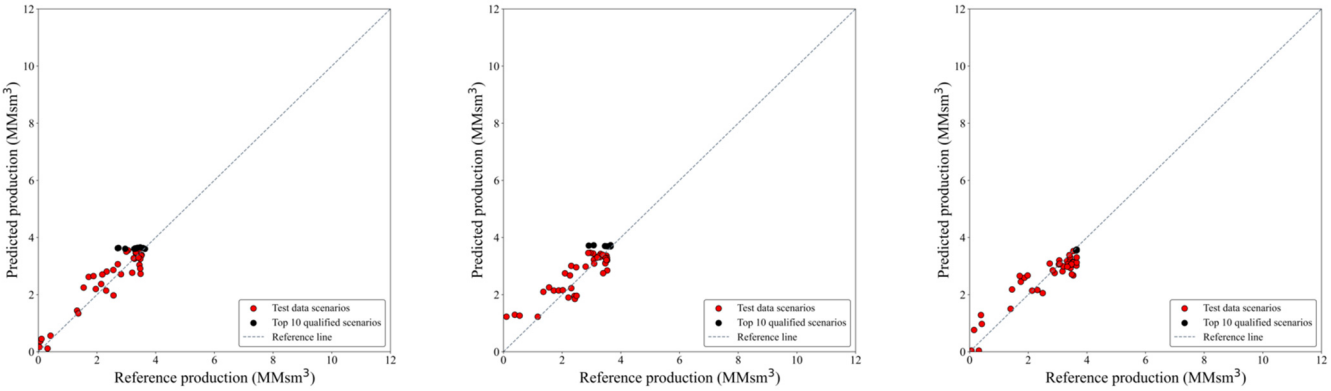
Table 4. Iterative learning performance of the M-CNN for the sequential well placement problem of the UNISIM-I-D.

Well	Number of Iterative Learning	Number of Feasible Well Locations	Number of Learning Data	R^2		ε_{avg} (%) for Ten Qualified Scenarios
				Training Data	Test Data	
1st well	Basecase	2282	176	0.93	0.87	13.34
	1st		186	0.85	0.77	7.13
	2nd (final)		196	0.89	0.84	2.64
2nd well	Basecase	2281	209	0.95	0.93	5.31
	1st (final)		219	0.94	0.90	1.66
3rd well	Basecase	2280	180	0.91	0.87	21.53
	1st (final)		190	0.96	0.94	0.68
4th well	Basecase (final)	2279	80	0.94	0.86	2.23

Second, compared with previous studies in which learning data are collected randomly, the preparation of learning data based on evolution was more efficient. This ensured the specialization of the constructed machine learning model for evaluating highly prolific regions. The number of learning data for each well placement was approximately 200, which was smaller than that reported by [26–28] (i.e., 504, 2000, and 400, respectively), although their target reservoirs were different from each other. The reduction in the cost of learning data preparation was further highlighted when the proposed approach was applied to field-scale simulations.

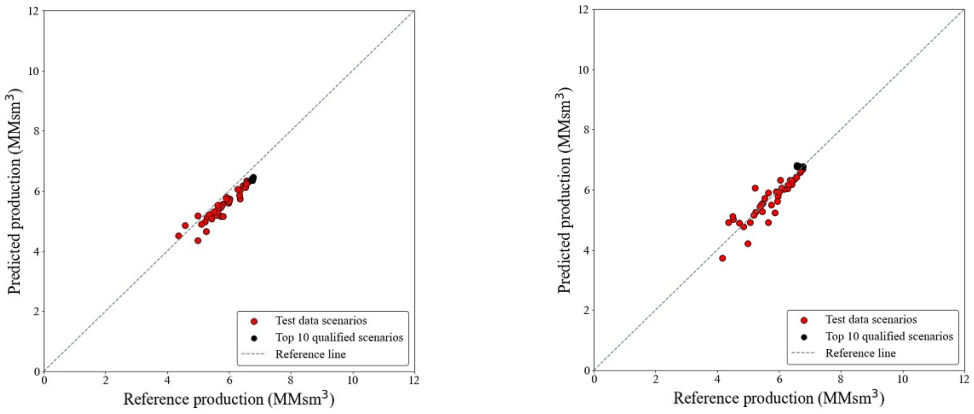
Third, the M-CNN with iterative learning exhibited reduced ε_{avg} compared to that without iterative learning owing to adding more high-productivity regions in its learning data. This contributes to the search for an optimal well location. The performance of the M-CNN in high-productivity regions increased with an increase in the number of iterative learning attempts if the computational cost was available. The behavior of the model within the ranges spanned by the learning data barely affects its generalization performance as a high-productivity region (i.e., N_{qual}) outside the learning data [45]. These results support the ability of iterative learning to overcome the intrinsic limitation of the extrapolation of neural networks by adding high Q_o^{NN} scenarios to the learning dataset for the next proxy implementation. Completing the iteration process implies that the extrapolation problem (i.e., the maximization problem) is almost entirely converted into an interpolation problem.

Figures 9–12 show the scatter plots used to compare the reference and predicted production for the test data scenarios and the top ten qualified scenarios for each well placement problem. In each subfigure, the red dots represent the Q_o^{RS} and Q_o^{NN} pairs for the test data. Similarly, the black dots represent pairs of the top ten scenarios. For the basecases, the proxies tended to overestimate the oil productivity for the top ten scenarios. As previously mentioned, this overestimation reveals the intrinsic limitation of machine learning prediction for extrapolation. This limitation could be overcome by implementing iterative learning. Based on these results, it is recommended to conduct iterative learning up to a level at which the improvement in evaluation performance becomes insignificant when applying machine learning for extrapolation, such as in well placement problems. Because the production wells were installed sequentially, the number of iterative learning applications decreased. In particular, the fourth well obtained high-quality results without iterative learning (Figure 12) because the M-CNN had already learned the input data with sufficient information.



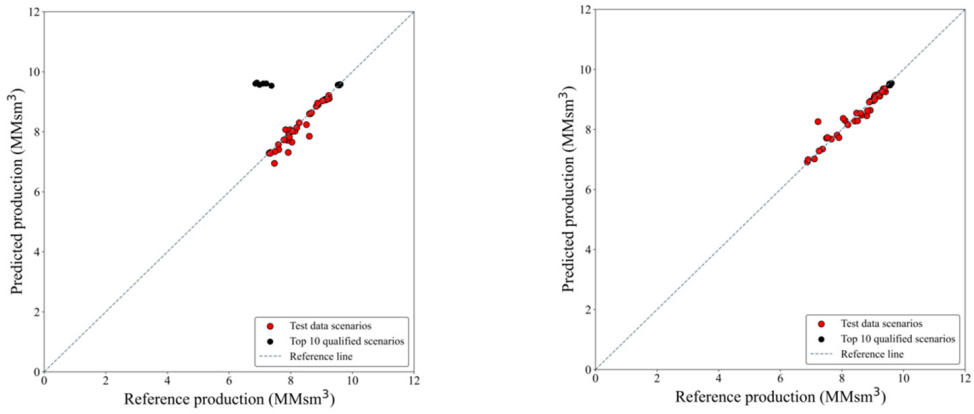
(**a**) Basecase (**b**) 1st iterative learning (**c**) 2nd iterative learning

Figure 9. Scatter plots to compare the reference and predicted production for test data scenarios and top ten qualified scenarios for the 1st well placement problem. (**a**) Basecase ($R^2 = 0.87$, $\varepsilon_{avg} = 13.34\%$); (**b**) 1st iterative learning ($R^2 = 0.77$, $\varepsilon_{avg} = 7.13\%$); (**c**) 2nd iterative learning ($R^2 = 0.84$, $\varepsilon_{avg} = 2.64\%$).



(**a**) Basecase (**b**) 1st iterative learning

Figure 10. Scatter plots to compare reference and predicted production for test data scenarios and top ten qualified scenarios for the 2nd well placement problem. (**a**) Basecase ($R^2 = 0.93$, $\varepsilon_{avg} = 5.31\%$); (**b**) 1st iterative learning ($R^2 = 0.90$, $\varepsilon_{avg} = 1.66\%$).



(**a**) Basecase (**b**) 1st iterative learning

Figure 11. Scatter plots to compare reference and predicted production for test data scenarios and top ten qualified scenarios for the 3rd well placement problem. (**a**) Basecase ($R^2 = 0.87$, $\varepsilon_{avg} = 21.53\%$); (**b**) 1st iterative learning ($R^2 = 0.94$, $\varepsilon_{avg} = 0.68\%$).

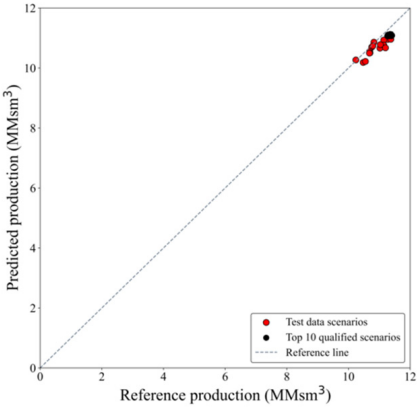
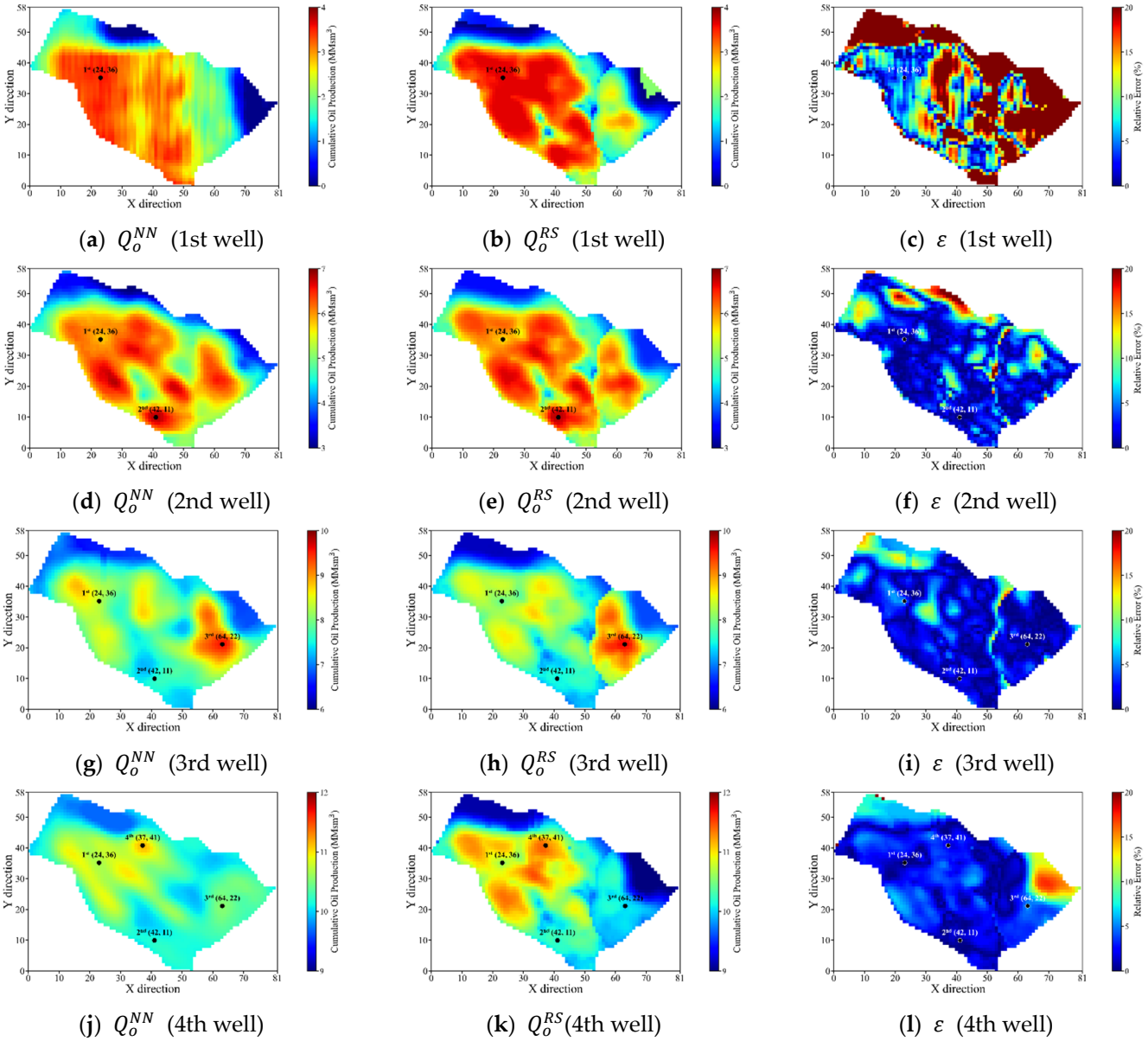


Figure 12. Scatter plots to compare reference and predicted production for test data scenarios and top ten qualified scenarios for the basecase of the 4th well placement problem ($R^2 = 0.86$, $\varepsilon_{avg} = 2.23\%$).

Figure 13 shows the oil productivity maps of the field cumulative oil production (i.e., Q_o^{NN} and Q_o^{RS} maps) obtained using M-CNN and RS. For quantitative comparison, relative error maps (i.e., ε maps) were drawn using Equation (8) (Figure 13c,f,i,l): In Figure 13, the black dots indicate the location of the well selected using the proposed approach. For most real-world problems, the acquisition of Q_o^{RS} maps is extremely difficult because of high computational costs (Figure 13b,e,h,k). By contrast, acquiring Q_o^{NN} maps is affordable and reliable, provided that reasonable accuracy is guaranteed for the test data and final solutions (Figure 13a,d,g,j). Overall, the Q_o^{NN} maps are similar to the corresponding Q_o^{RS} maps, particularly in highly prolific regions. Some large ε values were obtained along the reservoir boundary because of the low-proliferation regions, where the Q_o^{RS} values were significantly smaller than the Q_o^{NN} values. Owing to this overestimation of Q_o^{NN}, ε was higher than 20% because Q_o^{RS} is the denominator of Equation (8). Nevertheless, these low-proliferation regions were not the targets of this study. Instead, this study focused on the search for highly prolific regions. To achieve this, masking is useful in excluding lowly productive regions based on expert knowledge, either a priori or a posteriori [29]. Furthermore, the overall quality improved during the sequential well placement problem, as shown in Figure 13c,f,i,l.

Figures 14–16 show the production profiles at the four well locations selected using the M-CNN, PSO with ES, and the original study [38], respectively. To maintain the oil production rate close to the maximum surface liquid rate of 2000 m^3/day for as long as possible, the most significant factor affecting oil production was a minimum bottom-hole pressure of 150 kg/cm^2. Compared with the bottom-hole pressure, the water production was relatively insignificant in every well during primary production.

Table 5 compares the final well locations and corresponding field cumulative oil production: the production results of the original study [38], PSO results with ES, M-CNN results, and PSO for simultaneous well placement. These four methods yielded different well locations and drilling sequences, thereby yielding different four-year field cumulative productions. Both PSO and M-CNN recommended grid block coordinates (64, 22) as the drilling location of the third production well, indicating that both PSO and the proxy are useful in the optimal search. The proposed proxy exhibited 47.40% in field cumulative production compared with the original study [38]. Comparing the PSO with ES and PSO for simultaneous optimization, the improvement of the field cumulative oil production obtained using the M-CNN was 0.092 $MMsm^3$ (0.81%) and 0.033 $MMsm^3$ (0.29%), yielding an additional profit of 57.87 MMUSD and 20.75 MMUSD when the oil price is USD 100/STB, respectively. In summary, the accuracy of the developed proxy integrating PSO, M-CNN, and iterative learning in the decision making of well placement problems was validated.



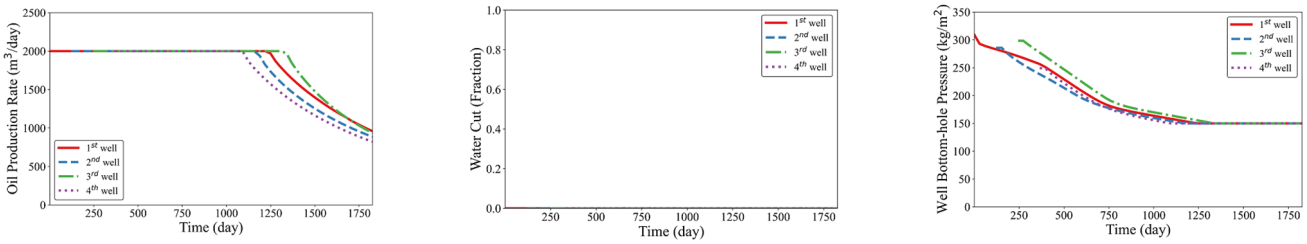
(**a**) Q_o^{NN} (1st well) (**b**) Q_o^{RS} (1st well) (**c**) ε (1st well)

(**d**) Q_o^{NN} (2nd well) (**e**) Q_o^{RS} (2nd well) (**f**) ε (2nd well)

(**g**) Q_o^{NN} (3rd well) (**h**) Q_o^{RS} (3rd well) (**i**) ε (3rd well)

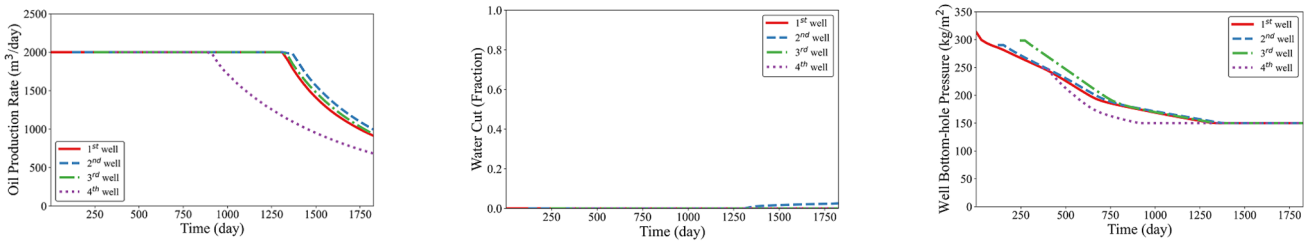
(**j**) Q_o^{NN} (4th well) (**k**) Q_o^{RS}(4th well) (**l**) ε (4th well)

Figure 13. Oil productivity and relative error (ε) maps obtained using M-CNN and reservoir simulation (RS). (**a**) Q_o^{NN} (1st well); (**b**) Q_o^{RS} (1st well); (**c**) ε (1st well); (**d**) Q_o^{NN} (2nd well); (**e**) Q_o^{RS} (2nd well); (**f**) ε (2nd well); (**g**) Q_o^{NN} (3rd well); (**h**) Q_o^{RS} (3rd well); (**i**) ε (3rd well); (**j**) Q_o^{NN} (4th well); (**k**) Q_o^{RS} (4th well); (**l**) ε (4th well).



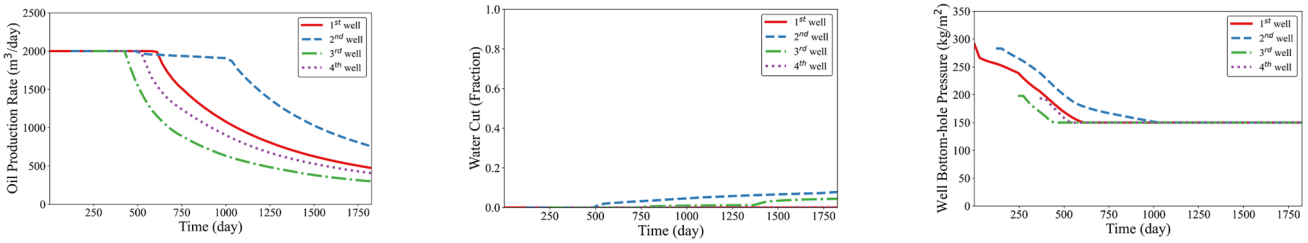
(**a**) Oil production (**b**) Water cut (**c**) Bottom-hole pressure

Figure 14. Well production profiles at the selected four well locations based on the M-CNN results. (**a**) Oil production; (**b**) water cut; (**c**) bottom-hole pressure.



(**a**) Oil production (**b**) Water cut (**c**) Bottom-hole pressure

Figure 15. Well production profiles at the selected four well locations based on the PSO with early stopping (ES) results. (**a**) Oil production; (**b**) water cut; (**c**) bottom-hole pressure.



(**a**) Oil production (**b**) Water cut (**c**) Bottom-hole pressure

Figure 16. Well production profiles at the four original well locations in the original study from the original study [38]. (**a**) Oil production; (**b**) water cut; (**c**) bottom-hole pressure.

Table 5. Comparison of the oil productivity with four well locations and their field cumulative oil production: the results of the original study [38], PSO results with early stopping (ES), M-CNN results, and PSO without ES for simultaneous well placement. Note that M-CNN was designed based on the PSO results with ES for an efficient generation of learning data. Also, PSO for simultaneous well placement failed to activate ES in this study.

Method		Final Well Location (X, Y)				Field Cumulative Oil Production (MMsm^3)
		1st Well	2nd Well	3rd Well	4th Well	
Original well placement [38]		(38, 36)	(21, 36)	(44, 43)	(31, 27)	7.742
Sequential well placement	PSO w/ES	(25, 24)	(24, 41)	(64, 22)	(46, 10)	11.320
	M-CNN	(24, 36)	(42, 11)	(64, 22)	(37, 41)	11.412
Simultaneous well placement	PSO w/o ES	(43, 13)	(63, 23)	(34, 34)	(14, 43)	11.379

Table 6 presents a comparison of computational costs based on the number of RS runs. The RS was compared for all feasible scenarios: PSO with ES, M-CNN, and PSO for simultaneous well placement. For sequential well placement, the PSO results with ES were obtained from $N_{pop} \times N_{es}$ RS runs. The cost of the M-CNN was the sum of the PSO with ES cost $N_{pop} \times N_{es}$ and the number of additional RS for iterative learning $N_{iter} \times N_{qual}$, where N_{iter} refers to the number of iterations. In the case of the PSO used for simultaneous optimization, the experimental settings were N_{pop} = 40 and N_e = 40, with early stopping. The ES was not activated in this case study because the particles maintained the evolution slowly. Thus, the number of RS runs reached 1600. Evidently, the M-CNN proxy exhibited the advantage of computational efficiency over the full search using the full-physics RS. The ratio of the computational cost of the proxy to that of the full search using the RS was 11.18% (=1020/9122 × 100%). Compared with the PSO with ES results, the M-CNN delivered a superior solution by increasing the computational cost by 4.08% (=(1020−980)/980 × 100%). The M-CNN results showed a decrease in computational cost of 36.25% (=(1600−1020)/1600 × 100%) compared with PSO with ES and simultaneous well placement. Similarly to Table 6, Table 7 summarizes the total time required for all

computations. The computational cost to execute the M-CNN was 541.5 (=271.0 + 129.2 + 101.8 + 39.5) seconds, which is only 1.6% (=541.5/(9957 + 10,288.9 + 9625.1 +3982.8) × 100%) of the PSO with ES cost.

Table 6. Comparison of the computational cost (the number of full-physics reservoir simulation (RS) runs): RS runs for all feasible scenarios, PSO with ES, M-CNN, and PSO without ES for simultaneous well placement.

Method		Number of RS Runs				
		1st Well (A)	2nd Well (B)	3rd Well (C)	4th Well (D)	Total (E = A + B + C + D)
RS for all feasible scenarios	Full search	2282	2281	2280	2279	9122
Sequential placement	PSO w/ES	280	300	280	120	980
	M-CNN	300	310	290	120	1020
Simultaneous placement	PSO w/o ES	1600				1600

Table 7. Comparison of the computational cost (time): RS runs for all feasible scenarios, PSO with ES, M-CNN, and PSO without ES for simultaneous well placement.

Method		Computational Cost for RS Runs (s)				Computational Cost for M-CNN (s)				Total Computational Cost (s)
		1st Well	2nd Well	3rd Well	4th Well	1st Well	2nd Well	3rd Well	4th Well	
RS for all feasible scenarios	Full search	75,739.6	75,706.4	75,673.2	75,640.0	-	-	-	-	302,759.2
Sequential placement	PSO w/ES	9293.2	9957.0	9293.2	3982.8	-	-	-	-	32,526.2
	M-CNN	9957.0	10,288.9	9625.1	3982.8	271.0	129.2	101.8	39.5	34,395.3
Simultaneous placement	PSO w/o ES	65,097.0				-	-	-	-	65,097.0

3.5. Discussion

Based on these numerical results, this study demonstrated that the integration of evolutionary optimization, machine learning, and iterative learning produced effective solutions for well placement optimization problems at a reasonable computational cost. While our focus was on optimizing well locations within a sequential well placement framework, future research aims to extend the methodology to incorporate well operating conditions (e.g., water injection rates during waterflooding, CO_2 injection rates for enhanced oil recovery or geological carbon storage). This expansion will enable optimization of both the spatial arrangement of wells and operational parameters, potentially enhancing overall field development strategies. For the sake of clarity and scope management, the current study primarily validated the performance of the proposed methodology for well location optimization. Subsequent research will address rate optimization and other operational considerations, building upon this foundational work. Furthermore, the framework can be applied to solve a variety of geoenergy engineering problems such as history matching [46,47], pipeline management [48,49], and reservoir-well-surface facility-integrated system design [50].

While this study demonstrates promising results, there are several avenues for improvement. Although PSO has proven effective and computationally efficient for well placement optimization, more advanced algorithms, such as the Artificial Hummingbird Algorithm (AHA), could yield superior convergence speed and accuracy [51]. Future work will explore integrating these advanced algorithms into our hybrid framework to

enhance optimization processes and performance in large-scale applications. Despite PSO's ability to reduce computational costs compared to exhaustive simulations, its role in dataset generation can become a bottleneck in large-scale optimization problems, particularly with heterogeneous or complex reservoir models. In such scenarios, PSO's initial population and convergence speed can significantly impact efficiency, potentially leading to extended computation times.

Furthermore, in large-scale reservoir models with numerous candidate well placements and high spatial resolution, the iterative learning process can become computationally intensive, especially when coupled with full-physics reservoir simulations. As the number of wells increases, so does the computational burden, making optimizing the balance between accuracy and efficiency crucial. Consequently, additional research is necessary to address scalability challenges in large field-scale applications. Our future studies will investigate the practicability of the algorithm for various complex problems, including enhanced oil recovery and CO_2 sequestration. These investigations will further refine our approach and extend its applicability to a broader range of reservoir management scenarios.

4. Conclusions

This study designed an M-CNN, learning the PSO results, and enhanced its performance with iterative learning to solve the sequential well placement problem in UNISIM-I-D. After running the RS, the qualified locations selected using the M-CNN integrated with PSO and iterative learning for the four wells were validated. In addition, the performance of the M-CNN was compared with that of PSO. The M-CNN constructed with the learning dataset obtained using an evolutionary algorithm (i.e., PSO) exhibited increased accuracy and reduced computational cost. Iterative learning enabled the intensive search of high-oil-potential regions, thereby overcoming the limitations of neural networks in an extrapolation problem. Using the developed hybrid method, the computational cost spent to evaluate all feasible well location cases was approximately 11.18% of the cost required for exhaustive evaluation through RS. Despite the constraints of limited training data, the M-CNN model, when integrated with PSO and iterative learning, demonstrated superior performance compared to PSO alone. This hybrid approach achieved remarkable accuracy with a relative error of less than 3% and significantly boosted cumulative oil production by 47.40% relative to the original well placement strategy. Notably, the M-CNN's predictive capabilities rivaled those of the exhaustive RS-based search, underscoring its efficiency and effectiveness. This achievement highlights the robustness of the M-CNN framework in optimizing well placement, even under data scarcity, and its potential to revolutionize reservoir management practices by combining high accuracy with computational efficiency. Another advantage of the proposed approach was its ability to swiftly test every candidate location, thereby providing a detailed quality map for further examination. These results demonstrate the efficacy of M-CNN for the multi-well placement optimization problem, as it achieved a balance between accuracy and computational cost savings. This rapid and precise approach can enhance decision making for well placement in multi-million-dollar geoenergy projects.

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

Funding: This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1A6A1A08025520) and the Ministry of Science and ICT (No. RS-2023-00208297). Additionally, this work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (No. RS-2022-00143644).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We appreciate Computer Modelling Group Ltd. (CMG) for providing their reservoir simulation software package. Baehyun Min and Minsoo Ji also express our gratitude to the Weather and Climate Data Convergence Analysis Specialized Graduate School supported by the Korea Meteorological Institute.

Conflicts of Interest: Author Minsoo Ji is currently employed by the company POSCO International. The remaining author 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. Zandvliet, M.J.; Handels, M.; van Essen, G.M.; Brouwer, D.R.; Jansen, J.D. Adjoint-based well placement optimization under production constraints. *SPE J.* **2008**, *13*, 392–399. [CrossRef]
2. Pouladi, B.; Karkevandi-Talkhooncheh, A.; Sharifi, M.; Gerami, S.; Nourmohammad, A.; VahidiLi, A. Enhancement of SPSA algorithm performance using reservoir quality maps: Application to coupled well placement and control optimization problems. *J. Petrol. Sci. Eng.* **2020**, *189*, 106984. [CrossRef]
3. Boah, A.E.; Kondo, O.K.S.; Borsah, A.A.; Brantson, E.T. Critical evaluation of infill well placement and optimization of well spacing using the particle swarm algorithm. *J. Pet. Explor. Prod. Technol.* **2019**, *9*, 3113–3133. [CrossRef]
4. Ding, S.; Lu, R.; Xi, Y.; Liu, G.; Ma, J. Efficient well placement optimization coupling hybrid objective function with particle swarm optimization algorithm. *Appl. Soft Comput.* **2020**, *95*, 106511. [CrossRef]
5. Wang, X.; Haynes, R.D.; Feng, Q. Well control optimization using derivative-free algorithms and a multiscale approach. *Comput. Chem. Eng.* **2019**, *123*, 12–33. [CrossRef]
6. Bertini, J.R., Jr.; Funcia, M.A.; Santos, A.A.; Schiozer, D.J. A comparison of machine learning algorithms as surrogate model for net present value prediction from wells arrangement data. In Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 14–19 July 2019.
7. Lu, R.; Reynolds, A.C. Joint optimization of well locations, types, drilling order and controls given a set of potential drilling paths. In Proceedings of the SPE Reservoir Simulation Conference, Galveston, TX, USA, 10–11 April 2019.
8. Wang, N.; Chang, H.; Zhang, D.; Xue, L.; Chen, Y. Efficient well placement optimization based on theory-guided convolutional neural network. *J. Petrol. Sci. Eng.* **2022**, *208*, 109545. [CrossRef]
9. Fonseca, R.R.-M.; Chen, B.; Jansen, J.D.; Reynolds, A. A stochastic simplex approximate gradient for optimization under uncertainty. *Int. J. Numer. Meth. Eng.* **2017**, *109*, 1756–1776. [CrossRef]
10. Kennedy, J.; Eberhart, R.C. A new optimizer using particle swarm theory. In Proceedings of the 6th International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995.
11. Lee, K.Y.; Park, J.B. Application of particle swarm optimization to economic dispatch problem: Advantages and disadvantages. In Proceedings of the 2006 IEEE PES Pewer Systems Conference and Esposition, Atlanta, GA, USA, 7 February 2006.
12. Lee, S.; Min, B.; Wheeler, M.F. Optimal design of hydraulic fracturing in porous media using the phase field fracture model coupled with genetic algorithm. *Comput. Geosci.* **2018**, *22*, 833–849. [CrossRef]
13. Schlumberger. *MEPO User Guide (Version 2020.01)*; Schlumberger: Houston, TX, USA, 2020.
14. CMG. *CMOST-AI User Guide (Version 2021.10)*; Computer Modeling Group LTD: Calgary, AB, Canada, 2021.
15. Salehian, M.; Haghighat, M.; Muradov, K. Multi-solution well placement optimization using ensemble learning of surrogate models. *J. Petrol. Sci. Eng.* **2022**, *210*, 110076. [CrossRef]
16. Min, B.; Sun, A.Y.; Wheeler, M.F.; Jeong, H. Utilization of multiobjective optimization for pulse testing dataset from a CO_2-EOR/sequestration field. *J. Petrol. Sci. Eng.* **2018**, *170*, 244–266. [CrossRef]
17. Guo, Z.; Reynolds, A.C.; Zhao, H. Waterflooding optimization with the INISIM-FT data-driven model. *Comput. Geosci.* **2018**, *22*, 745–761. [CrossRef]
18. de Brito, D.U.; Durlofsky, L.J. Well control optimization using a two-step surrogate treatment. *J. Petrol. Sci. Eng.* **2020**, *187*, 106565. [CrossRef]

19. Kuang, L.; Liu, H.; Ren, Y.; Luo, K.; Shi, M.; Su, J.; Li, X. Application and development trend of artificial intelligence in petroleum exploration and development. *Petrol. Explor. Dev.* **2021**, *48*, 1–14. [CrossRef]
20. Ali Ahmadi, M.; Bahadori, A. A LSSVM approach for determining well placement and conning phenomena in horizontal wells. *Fuel* **2015**, *153*, 276–283. [CrossRef]
21. Panja, P.; Velasco, R.; Pathak, M.; Deo, M. Application of artificial intelligence to forecast hydrocarbon production from shales. *Petroleum* **2018**, *4*, 75–89. [CrossRef]
22. Nwachukwu, A.; Jeong, H.; Pyrcz, M.; Lake, L.W. Fast evaluation of well placements in heterogeneous reservoir models using machine learning. *J. Petrol. Sci. Eng.* **2018**, *163*, 463–475. [CrossRef]
23. Wang, M.; Hui, G.; Pang, Y.; Wang, S.; Chen, S. Optimization of machine learning approaches for shale gas production forecast. *Geoenergy Sci. Eng.* **2023**, *226*, 211719. [CrossRef]
24. Zhu, Q.; Lin, B.; Yang, G.; Wang, L.; Chen, M. Intelligent production optimization method for a low pressure and low productivity shale gas well. *Petrol. Explor. Dev.* **2022**, *49*, 886–894. [CrossRef]
25. Centilmen, A.; Ertekin, T.; Grader, A.S. Applications of neural networks in multiwell field development. In Proceedings of the SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 3–6 October 1999.
26. Min, B.; Park, C.; Kang, J.M.; Park, H.J.; Jang, I.S. Optimal well placement based on artificial neural network incorporating the productivity potential. *Energy Sources* **2011**, *33*, 1726–1738. [CrossRef]
27. Jang, I.S.; Oh, S.E.; Kim, Y.; Park, C.; Kang, H. Well-placement optimisation using sequential artificial neural networks. *Energy Explor. Exploit.* **2018**, *36*, 433–449. [CrossRef]
28. Kwon, S.; Park, G.; Jang, Y.; Cho, J.; Chu, M.; Min, B. Determination of oil well placement using convolutional neural network coupled with robust optimization under geological uncertainty. *J. Petrol. Sci. Eng.* **2021**, *201*, 108118. [CrossRef]
29. Chu M -g Min, B.; Kwon, S.; Park, G.; Kim, S.; Huy, N.X. Determination of an infill well placement using a data-driven multi-modal convolutional neural network. *J. Petrol. Sci. Eng.* **2020**, *195*, 106805.
30. Kim, J.; Lee, K.; Choe, J. Efficient and robust optimization for well patterns using a PSO algorithm with a CNN-based proxy model. *J. Petrol. Sci. Eng.* **2021**, *207*, 109088. [CrossRef]
31. Atrey, P.; Hossain, A.; El Saddik, A.; Kankanhalli, M. Multi-modal fusion for multimedia analysis: A survey. *Multimed. Syst.* **2010**, *16*, 345–379. [CrossRef]
32. Dodge, S.; Karam, L. Understanding how image quality affects deep neural networks. In Proceedings of the 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 6–8 June 2016.
33. Busemeyer, J.R.; Byun, E.; Delosh, E.L.; McDaniel, M.A. *Learning Functional Relations Based on Experience with Input-Output Pairs by Humans and Artificial Neural Networks*; The MIT Press: Cambridge, MA, USA, 1997.
34. Delosh, E.L.; Busemeyer, J.R.; McDaniel, M.A. Extrapolation: The sine qua non for abstraction in function learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **1997**, *23*, 968–986. [CrossRef] [PubMed]
35. McDaniel, M.A.; Busemeyer, J.R. The conceptual basis of function learning and extrapolation: Comparison of rule-based and associative-based models. *Psychon. Bull. Rev.* **2005**, *12*, 24–42. [CrossRef] [PubMed]
36. Clerc, M.; Kennedy, J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Trans. Evol. Comput.* **2002**, *6*, 58–73. [CrossRef]
37. da Cruz, P.S.; Horne, R.N.; Deutsch, C.V. The quality map: A tool for reservoir uncertainty quantification and decision making. *SPE Reserv. Eval. Eng.* **1999**, *7*, 6–14. [CrossRef]
38. Gaspar, A.T.F.S.; Avansi, G.D.; Santos, A.D.S.; Hohendorff Filho, J.C.V.; Schiozer, D.J. UNISIM-I-D: Benchmark studies for oil field development and production strategy selection. *Int. J. Model. Simul. Pet. Ind.* **2015**, *9*, 47–55.
39. Almeida, F.; Davolio, A.; Schiozer, D.J. Reducing uncertainties of reservoir properties in an automatized process coupled with geological modeling considering scalar and spatial uncertain attributes. *J. Petrol. Sci. Eng.* **2020**, *189*, 106993. [CrossRef]
40. Gaspar, A.T.F.S.; Santos, S.M.G.; Ferreira, C.J.; Davolio, A.; Schiozer, D.J. Estimation of the chance of success of a four-dimensional seismic project for a developed oil field. *Comput. Geosci.* **2022**, *26*, 463–475. [CrossRef]
41. Negash, B.M.; Tufa, L.D.; Ramasamy, M.; Awang, M.B. System identification-based proxy model of a reservoir under water injection. *Model. Simul. Eng.* **2017**, *2017*, 7645470. [CrossRef]
42. Avansi, G.D.; Schiozer, D.J. UNISIM-I: Synthetic model for reservoir development and management applications. *J. Model. Simul. Petrol. Ind.* **2015**, *9*, 21–30.
43. Glorot, X.; Bordes, A.; Bengio, Y. Deep sparse rectifier neural networks. In Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 11–13 April 2011.
44. Gómez, D.; Rojas, A. An empirical overview of the no free lunch theorem and its effect on real-world machine learning classification. *Neural Comput.* **2016**, *28*, 216–228. [CrossRef] [PubMed]
45. Balestriero, R.; Pesenti, J.; LeCun, Y. Learning in high dimension always amount to extrapolation. *arXiv* **2021**, arXiv:2100.09485.
46. Jo, S.; Jeong, H.; Min, B.; Park, C.; Kim, Y.; Kwon, S.; Sun, A. Efficient deep-learning-based history matching for fluvial channel reservoirs. *J. Petrol. Sci. Eng.* **2022**, *208 Pt E*, 109247. [CrossRef]

47. Al-Mudhafar, W.J.; Vo, T.H.; Wood, D.A.; Min, B. Stochastic lithofacies and petrophysical property modeling for fast history matching in heterogeneous clastic reservoir applications. *Sci. Rep.* **2024**, *14*, 22. [CrossRef] [PubMed]
48. Shaik, N.B.; Jongkittinarukorn, K.; Benjapolakul, W.; Bingi, K. A novel neural network-based framework to estimate oil and gas pipelines life with missing input parameters. *Sci. Rep.* **2024**, *14*, 4511. [CrossRef]
49. Shaik, N.B.; Benjapolakul, W.; Pedapati, S.R.; Bingi, K.; Le, N.T.; Asdornwised, W.; Chaitusaney, S. Recurrent neural network-based model for estimating the life condition of a dry gas pipeline. *Process Saf. Environ.* **2022**, *164*, 639–650. [CrossRef]
50. Choi, S.; Kim, T.-W.; Choi, B.; Min, B. Design of integrated CO_2 transportation-injection-storage system based on multi-objective optimization. *J. Korean Inst. Min. Energy Resour. Eng.* **2024**, *61*, 347–362. [CrossRef]
51. Zhao, W.; Wang, L.; Mirjalili, S. Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications. *Comput. Methods Appl. Mech. Eng.* **2022**, *388*, 114194. [CrossRef]

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.