Advances in Geo-Energy Research

Vol. 8, No. 1, p. 37-44, 2023

Original article

Physics-informed machine learning for solving partial differential equations in porous media

Liqun Shan[1,2], Chengqian Liu[2], Yanchang Liu[2]*, Yazhou Tu[1], Linyu Dong[3], Xiali Hei[1]

[1]*School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA 70503, USA*

[2]*School of Physical and Electrical Engineering, Northeast Petroleum University, Daqing 163318, P. R. China*

[3]*Retirement Management Center, Dagang Oilfield, Tianjin 300456, P. R. China*

Keywords:
Porous media
two-phase flow
Buckley-Leverett equation
physics-informed neural networks
recurrent neural network
attention mechanism

Cited as:
Shan, L., Liu, C., Liu, Y., Tu, Y., Dong, L., Hei, X. Physics-informed machine learning for solving partial differential equations in porous media. Advances in Geo-Energy Research, 2023, 8(1): 37-44. https://doi.org/10.46690/ager.2023.04.04

Abstract:
Physical phenomenon in nature is generally simulated by partial differential equations. Among different sorts of partial differential equations, the problem of two-phase flow in porous media has been paid intense attention. As a promising direction, physics-informed neural networks shed new light on the solution of partial differential equations. However, current physics-informed neural networks' ability to learn partial differential equations relies on adding artificial diffusion or using prior knowledge to increase the number of training points along the shock trajectory, or adaptive activation functions. To address these issues, this study proposes a physics-informed neural network with long short-term memory and attention mechanism, an ingenious method to solve the Buckley-Leverett partial differential equations representing two-phase flow in porous media. The designed network structure overcomes the dependency on artificial diffusion terms and enhances the importance of shallow features. The experimental results show that the proposed method is in good agreement with analytical solutions. Accurate approximations are shown even when encountering shock points in saturated fields of porous media. Furthermore, experiments show our innovative method outperforms existing traditional physics-informed machine learning approaches.

1. Introduction

Physical phenomena in nature are usually modeled by partial differential equations (PDEs). Because the fields of PDEs are infinite dimensional spaces, they are quite difficult to solve those PDEs. In the past several decades, to accurately structure physical processes, predict physical phenomena, and push breakthroughs in geotechnical engineering discoveries, numerous researchers employed numerical methods to simulate physical systems, that is to say, numerical simulations. Finite-dimensional approximations have been developed, such as the finite element method, finite volume method, finite difference method, etc. Solutions to high-dimensional PDEs usually imply huge matrices, resulting in a large computational cost. In this case, the numerical schemes require dividing the space into multiple small grid blocks. Therefore, numerical simulations are not feasible for real-time and many query scenarios with high computational requirements.

With the popularity of machine learning, it has been one of the ubiquitous methods for solving physical engineering problems. At present, these techniques have been demonstrated to be a promising approach in solving the aforementioned issues (Cai et al., 2021; Jin et al., 2021; Almajid and Abu-Al-Saud, 2022; Kemeth et al., 2022; Vinuesa and Brunton, 2022). Machine learning approaches can be classified into two methodologies: data-driven methods and physics-informed neural networks (PINN) methods. Data-driven approaches cannot achieve good performance without large amounts of data. As a matter of fact, reservoir engineering data are classified as small data set owing to the complicacy of physics laws and the unpredictability of fluid movement within porous media. For instance, considering that drilling is only a ref-

Yandy Scientific Press

*Corresponding author.
E-mail address: liqun.shan1@louisiana.edu (L. Shan); chengqian_liu0302@163.com (C. Liu); ychang_liu@nepu.edu.cn (Y. Liu); yazhou.tu1@louisiana.edu (Y. Tu); linyu_deng@163.com (L. Dong); xiali.hei@louisiana.edu (X. Hei).

Received November 9, 2022; revised December 5, 2022; accepted January 4, 2023; available online January 10, 2023.

erence point that is unevenly dispersed in a huge reservoir, information is scarce (Jin et al., 2021). Data-driven models can make predictions faster than numerical simulators. Data-driven techniques outperform numerical simulations when the training dataset is easily accessible and the underlying physics equations are unknown. But the data-driven approach does have certain drawbacks (Almajid and Abu-Al-Saud, 2022). The fundamental problem is that the complex model structure is very difficult to analyze and evaluate, which makes it poorly interpretable and lacks system flexibility. Cai et al. (2021) review flow physics-informed learning, integrating seamlessly data and mathematical models, and implementing them using PINNs.

PINN is a representative example of physics-informed machine learning (PIML) in solving PDEs (Raissi et al., 2019; Khoo et al., 2021; Cuomo et al., 2022). In PINNs, PDEs indicating physical laws are typically incorporated into the loss functions of machine learning models to constrain objective optimization. The main breakthrough of PINN is the insertion of an automatic differential residual network that encodes physical formulas (Raissi et al., 2019). The great advantage of PINNs as PDE solvers is that they can handle high-dimensional problems with highly complex boundaries since there is no need to create complex meshes. PINN-based methods allow simulations to be limited only by the underlying physics and require no input samples. In PINNs, Neural Networks (NN) can be viewed as a gradient-based approach to PDE optimization (Khoo et al., 2021). PINN forces NN to minimize PDE residuals. Furthermore, NNs can act as near-universal functions because of their potential for nonlinear prediction and data integration (Cuomo et al., 2022). They can estimate any invisible and unmeasurable phenomena with the required precision. However, it is essential to reassemble the spatially dynamic fluid to account for low-cost computation. Fukami et al. (2019) proposed a high-precision framework, which has been shown to be effective in generating turbulence from unusually rough flow fields. For more complex problems, more hidden units are needed, and more and more parameters need to be updated in the network, which involves additional computation.

It must be admitted that the PINN process is particularly time intensive to complicated PDE problems, lacking generalization and losing accuracy when working with complex grids. From this point of view, traditional PINN-based techniques are constrained by low-dimensional systems and are better at straightforwardly time-independent physics phenomena. PINNs have disadvantages in issues with drastic gradients or issues requiring coupled PDEs. To get rid of time-dependent issues, researchers employ continuous-time approaches, in detail, regarding time and spatial dimensions in the same way (Meng et al., 2020). The initial modality of PINNs has been improved by introducing domain fragmentation techniques (Jagtap and Karniadakis, 2020), nonlinear formulas (Haghighat et al., 2021), adjustable weighting on the optimization issue or novel network with feature embedding (Wang et al., 2021).

In recent years, deep learning has been introduced to study two-phase flow in porous media, which has attracted many academic interests due to its ability to deal with nonlinearity and high dimensions. Due to the high generalization quality of deep learning, it is widely considered to be a better solution for dealing with computational PDEs. Encoding the PDEs into the deep learning model essentially reduces the size of the training dataset required to adjust the parameters of deep learning models. PIML and physics-informed deep learning are active research topics (Wu et al., 2018; Raissi et al., 2019; Karniadakis et al., 2021; Lu et al., 2021; Ranade et al., 2021). Several recent studies have demonstrated the advantages of using deep learning models to solve Buckley-Leverett PDEs (Almajid and Abu-Al-Saud, 2022). Fuks and Tchelepi (2020) optimized a two-phase immiscible transport problem in porous media with the help of a second-order derivative term using PINN methods. They employ artificial dissipation to discover answers using shocked and blended waves. Rodriguez-Torrado et al. (2022) proposed informed attention-based neural networks (PIANNs). PIANNs combine Gated recurrent units and attention mechanisms to overcome the limitations of current PINNs by adding artificial dissipation. This work motivates us to explore the use of recurrent neural networks and attention mechanisms to provide high-quality solutions to Buckley-Leverett PDEs.

This study is inspired by existing works (Haghighat et al., 2021; Cuomo et al., 2022; Rodriguez-Torrado et al., 2022). In view of the original PINN structure, an innovative deep learning approach was developed to solve nonlinear partial differential equations in porous media by pairing a long short-term memory (LSTM) with an attention mechanism. Experiments were carried out on various NNs to examine their prediction performance, especially convolutional neural networks (Esmaeilzadeh et al., 2020), residual convolutional neural networks (Coutinho et al., 2022), residual networks (Cheng et al., 2021; Hanna et al., 2022), PINN approach (Fuks and Tchelepi, 2020), and generative adversarial network (Yang and Perdikaris, 2019). With tens of thousands of training and testing, the proposed method reaches more accurate solutions than existing methods. The remaining of this paper is the following. Section 2 describes an oil-water system in porous media and gives an account of the standard Buckley-Leverett model with oil and water. Section 3 illustrates the traditional PINN and our proposed approach, listing network structure and loss function in our proposed physics-informed LSTM with attention mechanism. Experiments and results are solved in Section 4, minutiae including experimental results and comparisons with classic PINNs. Section 5 discusses this paper's work. Conclusions are drawn in Section 6.

2. Oil-water system in porous media

This study considers the standard Buckley-Leverett model with oil and water. In a porous medium of permeability $k(x)$ and porosity $\phi(x)$, the wetting phase, water (w), displaces the nonwetting phase, oil (o). Gravity and capillary effects are neglected. The water and oil saturation are governed by a mass balance coupled system complimented by Darcy's law equations. According to Corey-Brooks relative permeabilities, the water saturation can be governed by a non-linear Buckley-

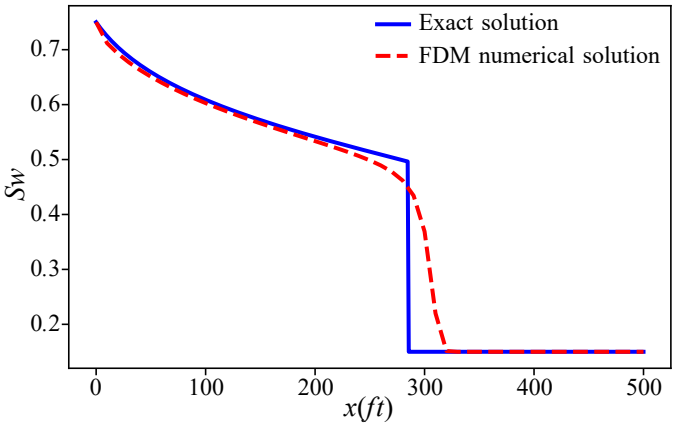

Fig. 1. The comparison between analytical solution and FDM solution.

Leverett equation:

$$\frac{\partial S(x,t)}{\partial t}+\frac{\mathrm{d}f(S(x,t))}{\mathrm{d}x}=0 \tag{1}$$

where $S(x,t)$ is the saturation of water. $f(\cdot)$ is the fractional flow function. $f(S(x,t))$ is the fractional flow of water. Independent variables x and t are dimensionless space and time and are expressed as:

$$t=\frac{q}{\phi AL}t_T, x=\frac{x_T}{L} \tag{2}$$

where q refers to the total flow rate, ϕ is the porosity, L denotes the length, A is the cross-sectional area, t_T is the dimensional time, and x_T represents the dimensional distance.

At the initial condition ($t=0$), the saturation of water is S_{wc}. At any time and at the boundary condition ($x=0$), the saturation of water is $1-S_{or}$. Boundary and initial conditions can be marked as:

$$S(x,t=0)=S_{wc}, S(x=0,t)=1-S_{or} \tag{3}$$

Specifically, the result of $S(x,t)$ is re-expressed as u. In this case, Eq. (1) is simplified as:

$$u_t+f(u)_x=0 \tag{4}$$

The dependent variable u is given by the network based on variables of space coordinate x, time coordinate t. The fractional flow function of water $f(S)$ in Eq. (1) is expressed as:

$$f(S)=\frac{M(S-S_{wc})^2}{M(S-S_{wc}^2)+(1-S-S_{or})^2} \tag{5}$$

where M is the mobility ratio, which is calculated as the proportion between oil viscosity (μ_o) and water viscosity (μ_w). M is assumed to be proportional to 2 in our study. The PDE in Eq. (1) and non-convex flux function $f(S)$ in Eq. (5) make up classical Buckley-Leverett problem that mentioned above.

The Fig. 1 shows the comparison between analytical solution and numerical solution by using FDM where time $t=15$ days. It can be seen that the numerical scheme fails to track the shock location. Next, it is necessary to correct the solution to better approximate the saturation of water.

3. PINN structure optimization in solving Buckley-Leverett problem

Researchers attempted to use PINNs to address discontinuous issues equipping non-convex flow function with shock points, which is known as the Buckley-Leverett equation (Fuks and Tchelepi, 2020; Fraces and Tchelepi, 2021). Fortunately, Buckley-Leverett model is the same to explore immiscible oil-water saturation in petroleum reservoir engineering (Lu et al., 2018). Fraces and Tchelepi (2021) provided solution to the Buckley-Leverett dilemma, in which the entropy and velocity restrictions are included in the NN residual. Although this approach can well capture the shock front, there is difficulty to reconstruction results. Xu et al. (2021) presented that the weak form formulation of the PDEs into the loss function to deal with strong discontinuities in two-phase flow issues. Unfortunately, the Buckley-Leverett equations used in the preceding two studies are not suitable to the realistic reservoir, they are only simplified equations that explain water saturation movement. The occurrence of discontinuities is the fundamental concern with using PINN to solve the Buckley-Leverett problem.

3.1 PINN

Previous works on solving the Buckley-Leverett equation use PINN structures. PINN's architecture has been almost exclusively based on fully connected feed-forward NNs. Using PINN's original framework (Raissi et al., 2019). Fuks and Tchelepi (2020) introduced an artificial dissipation term to PDE, which contribute to match the shock front. So, a second-order derivative term is added to the original Buckley-Leverett PDE. Eq. (1) is redefined as the following expression:

$$Z(x,t)=\frac{\partial S(x,t)}{\partial t}+\frac{\mathrm{d}f(S(x,t))}{\mathrm{d}x}-\varepsilon\frac{\partial^2 S(x,t)}{\partial x^2} \tag{6}$$

where ε is regarded as a coefficient parameter, generally, its order of magnitude is 10^{-2} or 10^{-3}. Fuks and Tchelepi (2020) set ε to 2.5×10^{-3}, PINN retrieves the proper solution at shock fronts.

The precondition of solving nonlinear PDE using PINN is selection of the compatible NN architecture, and corresponding hyperparameters also need be tweaked. Our approach follows the work of Raissi et al. (2019) for comparable situations. Fig. 2 demonstrates the simplified architecture of PINN for solving Buckley-Leverett PDE. In NN part, there are two parameters in the input layer, spatial coordinates x and time t. At the input layer x and t provided, the result of the output layer generation is water saturation, marked as u. NN with physics informed constraints are used to describe PDEs. Fuks and Tchelepi (2020) demonstrated that PINNs underperform to extract the best answer of the hyperbolic PDE. The study shows that when shocks are existent, the PIML methods are unsuitable for hyperbolic PDEs with discontinuity values.

3.2 The proposed LSTM-AM approach

Current PINNs' ability to learn PDEs relies on adding artificial diffusion or using prior knowledge to increase the

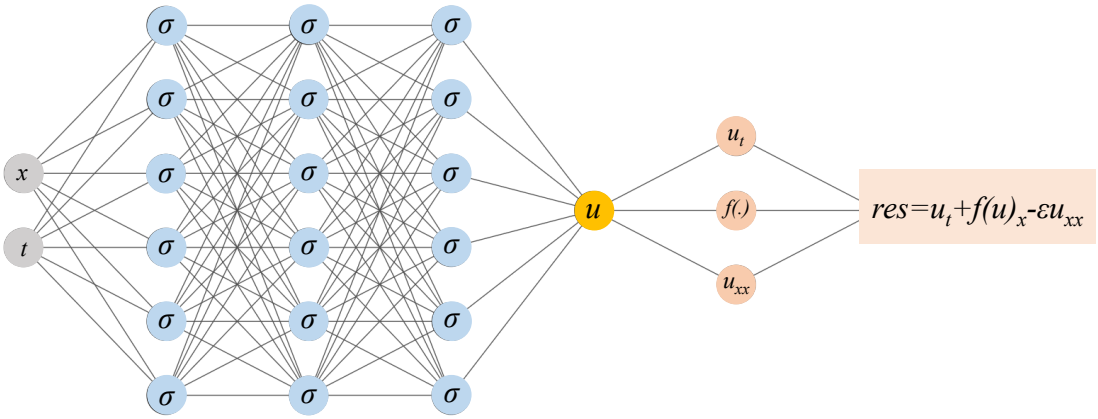

Fig. 2. The architecture of simplified PINN with PDE residual.

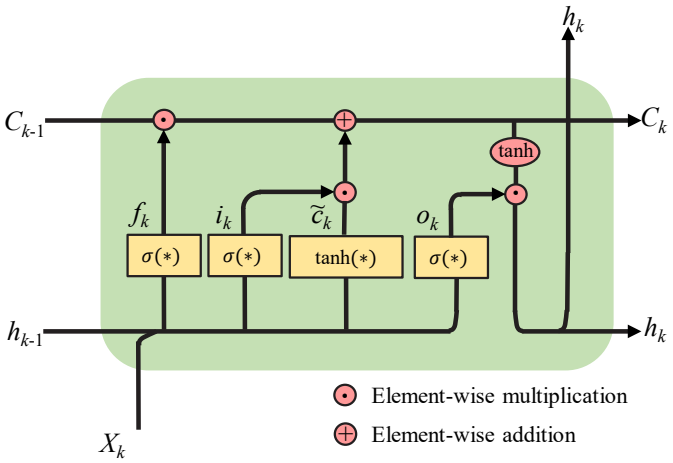

Fig. 3. A LSTM unit.

number of training points along the shock trajectory or adaptive activation functions. To address these issues, a framework based on LSTM and attention mechanism is designed to solve the Buckley-Leverett PDE.

LSTM is a powerful deep learning method to model sequence-to-sequence input-output relationship. It can better capture the dependency of long distance. It inherits the most characteristics of recurrent neural network and solves the problems of gradient disappearance/gradient explosion in the process of long sequence training.

A LSTM cell shown in Fig. 3 cosmists of a set of independent weights and biases that are shared across the entire temporal space within the cell. Each LSTM contains four interacting units, including an input gate i_k, an internal cell ($\tilde{c}_k$), a forget gate f_k, and an output gate o_k. The input gate guides the input flow into the internal cell state. The internal cell memorizes the cell states from the previous time step through self-loop connections. The output gate controls the output flow into the LSTM cell output. With the interaction of input/forget/output gates, cell states in the LSTM can selectively disseminate valuable information along temporal/spatial sequences to capture long short-term dependencies in a dynamical system.

The relationship among these gates can be described by the equations as follows:

$$i_k = \sigma(W_{xi}x_k + W_{hi}x_{k-1} + b_i) \tag{7}$$
$$f_k = \sigma(W_{xf}x_k + W_{hf}x_{k-1} + b_f) \tag{8}$$
$$\tilde{c}_k = \tanh(W_{xc}x_k + W_{hc}x_{k-1} + b_c) \tag{9}$$
$$o_k = \sigma(W_{xo}x_k + W_{ho}x_{k-1} + b_o) \tag{10}$$
$$c_k = f_k \odot c_{k-1} + i_k \odot \tilde{c}_k \tag{11}$$

where σ means the logistic sigmoid function. $W_{\alpha\beta}$ with $\alpha = \{x, h\}$ and $\beta = \{f, i, c, o\}$ refers to the weight matrices in terms of different inputs within different gates. b_β denotes the corresponding bias vectors. tanh denotes the hyperbolic tangent function. $\odot$ is the Hadamard product (element-wise product). As can be seen from the Fig. 5, hidden sequence $[h_1, h_2, \cdots, h_{k-1}]$ and cell sequence $[c_1, c_2, \cdots, c_{k-1}]$ assist in inputting sequence $[x_1, x_2, \cdots, x_k]$, which exports correspondingly $[h_2, \cdots, h_k]$ and $[c_2, \cdots, c_k]$, especially intermediate transition cell state $[\tilde{c_2}, \cdots, \tilde{c_k}]$. The complex connection mechanism makes a LSTM based network powerful in mapping the temporal/spatial long-term feature maps to the corresponding output space.

Attention mechanisms (AM) can allocate computing resources to more important tasks and filter out irrelevant information when computing power is limited (Vaswani et al., 2017). The attention mechanism is chosen because it has fast convergence speed, few parameter settings, and the ability of capturing significant features, which improves the efficiency and accuracy of task processing. After feature sequences go through hidden state h_t and cell state C_t in the LSTM unit, α_t^k is defined as a tangent activation function operation result of k^{th} input features at time t. Next, α_t^k is normalized with the help of a softmax function. β_t^k represents an attention weight, a which is the score of how much attention should be devoted on the k^{th} feature sequence. z_t is the output of the attention mechanism block at time t.

Inheriting LSTM advantages, assisted by attention mechanism, this study makes further efforts to pursue more nuance flow fluctuations in porous media. Borrowed the idea from (Wang et al., 2021; Rodriguez-Torrado et al., 2022), LSTM is integrated with attention for more precise prediction. In our study, LSTM with attention mechanism network structure cap-

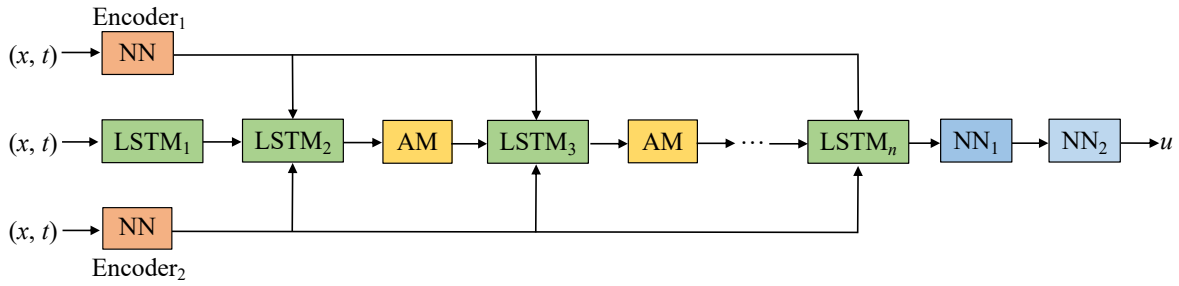

Fig. 4. The structure of the proposed LSTM-AM.

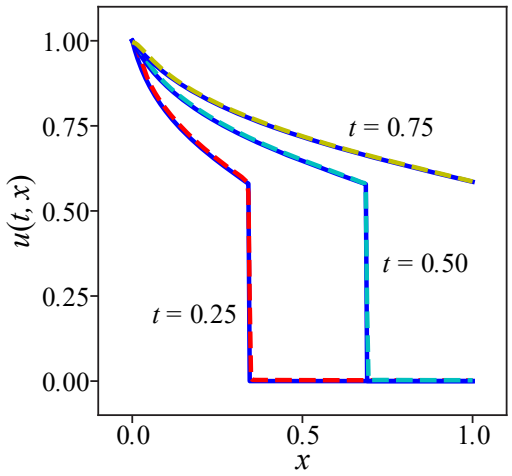

Fig. 5. Comparison of the predicted and exact solutions of Buckley-Leverett equation.

ture the location of the saturation front in this case. As shown in Fig. 6, the solutions obtained by our LSTM-AM method are more accurate than those obtained by CNN, ResCNN, NN, ResNN, and GAN at different time points at t = 0.25, 0.50, and 0.75.

(LSTM-AM) is regarded as function approximators. Fig. 4 depicts the architecture of solving Buckley-Leverett equation with our LSTM-AM.

The output of the network is marked as u. LSTM-AM can optimize $u(x,t)$ and PDE solution. For each PDE solution function u, the initial conditions (ICs) are described as u_0, $t = 0$. The boundary conditions (BCs) are determined for each spatial dimension depending on the geometrical issues. The hyperbolic tangent activation function and sigmoid function are applied in LSTM layer and attention layer, respectively.

The LSTM-AM can explicitly account for multiplicative interactions between different input dimensions and enhance the hidden states with residual connections. As shown in Fig. 4, the trunk of the network comprises LSTM and AM. Meanwhile, introduction of two encoders with conventional fully connected architectures maps the inputs variables to a high-dimensional feature space. and then use a point-wise multiplication operation to update the hidden layers according to the following forward propagation rule:

$$\text{Encoder}_1 = \tanh(\text{NN}(X)), \text{Encoder}_2 = \tanh(\text{NN}(X)) \quad (12)$$

$$H^1 = \text{LSTM}_1(X) \quad (13)$$

$$Z^1 = \tanh(H^1 \odot \text{Encoder}_1 + (1 - H^1) \odot \text{Encoder}_2) \quad (14)$$

$$H^1 = \text{LSTM}_l(\text{AM}(Z^{l-1})), l = 1, 2, \cdots, n-1 \quad (15)$$

$$Z^l = \tanh(H^l \odot \text{Encoder}_1 + (1 - H^l) \odot \text{Encoder}_2) \quad (16)$$

$$H^{l+1} = \text{LSTM}_{l+1}(\text{AM}(Z^l)) \quad (17)$$

$$u_\theta(X) = \tanh(\text{NH}_2(\tanh(\text{NH}_1(H^n)))) \quad (18)$$

3.3 Loss function

To constrain the PDE to the deep learning model, the model is enforced to satisfy the Eq. (1) by embedding Eq. (1) into the loss function. The training goal is to minimize the total regularized loss involving the collected data points from the PDE residual, initial conditions, boundary conditions, and observations. The various loss terms are defined as follows.

PDE residual loss:

$$\text{loss}_r = \frac{1}{N_R} \sum_i^{N_R} \left| \frac{A\phi}{q} \frac{\partial S}{\partial t} + \frac{\partial f(S)}{\partial x} \right|^2 \quad (19)$$

Boundary condition loss:

$$\text{loss}_b = \frac{1}{N_b} \sum_i^{N_b} |S_p - S_b|^2 \quad (20)$$

Initial condition loss:

$$\text{loss}_i = \frac{1}{N_i} \sum_i^{N_i} |S_p - S_i|^2 \quad (21)$$

where A is the cross-sectional area, q refers to the flow rate, and ϕ is the porous medium porosity. S_p represents the saturation value of the point in the x dimension at a certain time. N_r, N_b, and N_i are the number of collocated, boundary condition and initial condition points, respectively. For all experiments, the training data consist of $N_i = 250$ and $N_b = 250$ randomly distributed points on initial and boundary conditions and $N_r = 2,000$ collocation points for the residual term. Note that there are no observations in the training data, and a small dataset is used. All data points were sampled randomly over the interior of the domain $x \in [0, 1]$ and $t \in [0, 1]$. Regularizing the loss terms and summarize them, the total loss is obtained as:

$$\text{loss} = \lambda_r \text{loss}_r + \lambda_b \text{loss}_b + \lambda_i \text{loss}_i \quad (22)$$

where λ_r, λ_b and λ_i are optimized through experiments.

The workflow of our method is described as four steps. First, the deep learning model is built. Second, data points used in training, validation, and testing are collected from initial

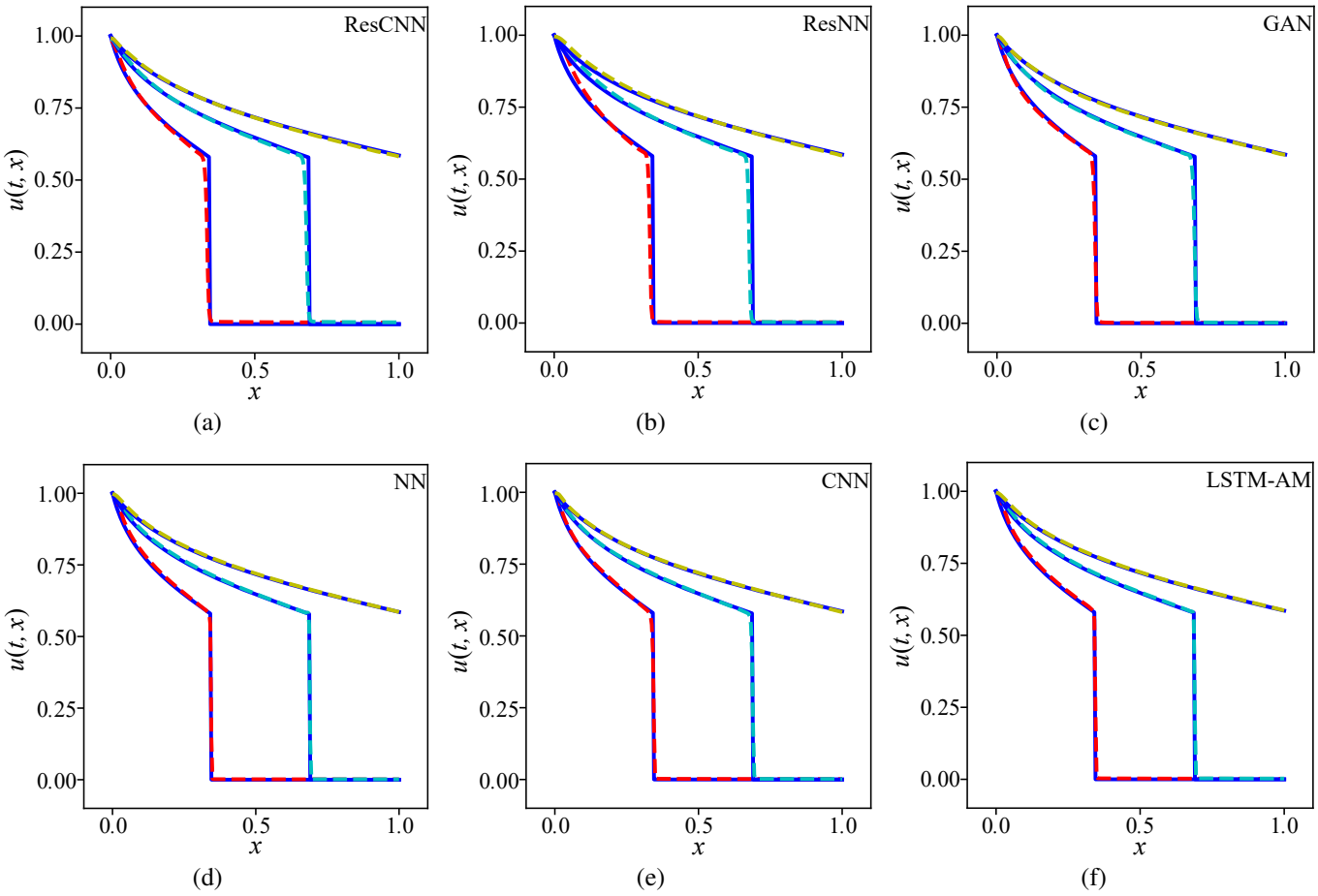

Fig. 6. The comparisons of the predictions from different algorithms.

condition, boundary condition, and collocation points. Third, a gradient-based optimization method (Adam optimizer) is applied to PDE for calculating the gradient descent direction of model's parameters. As the most popular optimization method, Adam optimizer keeps minimizing the loss value; it is feasible to lower the loss by an order of magnitude with tens of thousands more iterations. Forth, the deep learning model is calibrated and start next loop to continue the training for optimizing the parameters of model, proceeding with optimally approximating u with the help of our proposed LSTM-AM network structure. All weights are initialized randomly according to the Xavier initialization scheme (Glorot and Bengio, 2010).

4. Experiments

This section shows the results predicted by our model. To evaluate the accuracy of PINNs, the results from the PINNs models are compared with the analytical solutions of the Buckley-Leverett problem. Our method is also compared with other mainstream neural solver architectures.

4.1 Experimental results

Taking different dimensionless time points $t = 0.25$ (red dashed line), 0.50 (cyan-blue dashed line), and 0.75 (yellow dashed line) as an example, The LSTM-AM solution is shown in Fig. 5. It can be seen that the network successes in this case to provide a close approximation (dashed lines) of the underlying analytical solution (blue lines). Specially, the shock is well captured and nearly indiscernible from the exact PDE solution.

4.2 Comparisons with existing PINNs

Our LSTM-AM model is compared with other NN-based techniques, including convolution neural network (CNN) (Esmaeilzadeh et al., 2020), residual CNN (ResCNN) (Coutinho et al., 2022), NN (Fuks and Tchelepi, 2020), residual NN (ResNN) (Cheng et al., 2021), and generative adversarial network (GAN) (Yang and Perdikaris, 2019). All compared models use ICs, BCs, and PDE constraints.

Fig. 6 shows a comparison of the predicted solutions of PIMLs (dashed red line) with the exact solution (solid blue line) of the Buckley-Leverett equation at three different instants. Note that except our method, other PIML approaches used the diffusion term mentioned in Section 3.1 Eq. (6) to

Physics-informed learning identifies a number of parameters that not only fit the input data but also satisfy the physical phenomena in the manner of PDEs, initial conditions, and boundary conditions. This is accomplished by adding a suitable optimizer and minimizing loss function. Extending

Table 1. The comparisons of the final training loss from different algorithms.

Model	Loss
ResNN	8.909×10^{-5}
ResCNN	6.813×10^{-5}
GAN	5.617×10^{-5}
NN	1.512×10^{-5}
CNN	6.813×10^{-6}
LSTM-AM	7.051×10^{-7}

Table 2. The comparisons of the L_2 relative errors from different algorithms.

Model	L_2 relative errors
ResNN	0.08648
ResCNN	0.06806
GAN	0.05157
NN	0.03532
CNN	0.02096
LSTM-AM	0.01972

the number of trainings often improves solution accuracy but accompanying a computational expense. As a black box, NN needs to be trial and error tested, so we ran 20,000, iterations to train the PINNs. In the influence of network parameters and floating-point operations per second, as an effective counterweight, the optimal solution is found with different iterations. To capitalize on comparative experiments (Fuks and Tchelepi, 2020; Cheng et al., 2021; Fraces and Tchelepi, 2021; Haghighat et al., 2021), the proposed LSTM-AM successfully cut down final training loss to 10^{-7}, lower than others as shown in Table 1.

To further compare these methods with our approach, the L_2 relative error is computed for the entire solution. The L_2 relative error is expressed as the following:

$$L_2 \text{ relative error} = \frac{||u - \hat{u}||_2^2}{||u||_2^2} \tag{23}$$

where u is the real solution, $\hat{u}$ is the predictions obtained by the PINN methods. The smaller the relative error, the better the match between the true and the estimated solution. From the Table 2, it can be clearly seen that LSTM-AM's average L_2 relative error is lower than other models. This indicates the proposed method in this study has better accuracy in solving the Buckley-Leverett PDEs than methods of CNN, GAN, NN, ResCNN and ResNN.

5. Discussion

Once machine learning algorithms are equipped with the physics laws, it can predict PDE solutions more efficient than traditional numerical simulations. In this study, a novel PIML method is presented to address the issue of oil drainage in water-filled porous medium. Fig. 5 shows that our predictive solution is in perfect agreement with exact Buckley-Leverett solution. After training, the model offers appropriate reservoir simulations at each time-step. Unfortunately, this study concentrates on nonconvex flux functions in the PDE, ignoring convex condition. To consider the full range of flow simulations, the next step is to parameterize the convex hull of the flux function and further put the convex theory into practice. Furthermore, based on the experience summarized in (Alwated and El-Amin, 2021; Yang et al., 2021), further work will be done on multiscale simulation of multi-physics problems in porous media.

6. Conclusion

It is difficult to solve the partial differential equation of non-convex flux function composed of Buckley-Leverett problem in porous media flow. In this study, an innovative PINN with long-short-term memory and attention mechanism architecture is employed to solve the Buckley-Leverett flow problem, focusing on nonlinear two-phase transport in porous media. In the proposed method, the required prior knowledge comes from the initial/boundary conditions and the PDE. The underlying PDE is incorporated into the loss function of the neural network. The solution of the PDE can be predicted without any labeled data. Compared to the previous variant NN, our method provides an efficient and accurate method for solving partial differential equations in porous media. It can capture the fluid shock characteristics and ensures that the predicted output conforms to the relevant physical properties.

Acknowledgements

This research was financially supported by the Natural Science Foundation of Hebei Province (No. E2021107005), Northeast Petroleum University Foundation Founding (No. 2018GP2D-04). We gratefully acknowledge the helpful comments of the editor and anonymous reviewers.

Conflict of interest

The authors declare no competing interest.

References

Almajid, M. M., Abu-Al-Saud, M. O. Prediction of porous media fluid flow using physics informed neural networks. Journal of Petroleum Science and Engineering, 2022, 208: 109205.

Alwated, B., El-Amin, M. F. Enhanced oil recovery by nanoparticles flooding: From numerical modeling improvement to machine learning prediction. Advances in Geo-Energy Research, 2021, 5(3): 297-317.

Cai, S., Mao, Z., Wang, Z., et al. Physics-informed neural networks (PINNs) for fluid mechanics: A review. Acta

Mechanica Sinica, 2021, 37(12): 1727-1738.

Cheng, C., Zhang, G. T. Deep learning method based on physics informed neural network with resnet block for solving fluid flow problems. Water, 2021, 13(4): 423.

Coutinho, E. J. R., Dall'Aqua, M., McClenny, L., et al. Physics-informed neural networks with adaptive localized artificial viscosity. ArXiv preprint, 2022: 2203.08802.

Cuomo, S., Di Cola, V. S., Giampaolo, F., et al. Scientific machine learning through physics-informed neural networks: Where we are and what's next. Journal of Scientific Computing, 2022, 92(3): 88.

Esmaeilzadeh, S., Azizzadenesheli, K., Kashinath, K., et al. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. Paper 20489143 Presented in SC20: International Conference for High Performance Computing, Networking, Storage and Analysis, Virtual Event, 9-19 November, 2020.

Fraces, C. G., Tchelepi, H. Physics informed deep learning for flow and transport in porous media. Paper SPE 203934 Presented at the SPE Reservoir Simulation Conference, Virtual Event, 26 October-1 November, 2021.

Fukami, K., Fukagata, K., Taira, K. Super-resolution reconstruction of turbulent flows with machine learning. Journal of Fluid Mechanics, 2019, 870: 106-120.

Fuks, O., Tchelepi, H. A. Limitations of Physics Informed Machine Learning for Nonlinear Two-Phase Transport in Porous Media. Journal of Machine Learning for Modeling and Computing, 2020, 1(1): 19-37.

Glorot, X., Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. Paper Presented in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13-15 May, 2010.

Haghighat, E., Bekar, A. C., Madenci, E., et al. A nonlocal physics-informed deep learning framework using the peridynamic differential operator. Computer Methods in Applied Mechanics and Engineering, 2021, 385: 114012.

Hanna, J. M., Aguado, J. V., Comas-Cardona, S., et al. Residual-based adaptivity for two-phase flow simulation in porous media using physics-informed neural networks. Computer Methods in Applied Mechanics and Engineering, 2022, 396: 115100.

Jagtap, A. D., Karniadakis, G. E. Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. Paper Presented at AAAI Spring Symposium: MLPS, Stanford, CA, 22-24 March, 2021.

Jin, X., Cai, S., Li, H., et al. NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations. Journal of Computational Physics, 2021, 426: 109951.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., et al. Physics-informed machine learning. Nature Reviews Physics, 2021, 3(6): 422-440.

Kemeth, F. P., Bertalan, T., Thiem, T., et al. Learning emergent partial differential equations in a learned emergent space. Nature Communications, 2022, 13(1): 3318.

Khoo, Y., Lu, J., Ying, L. Solving parametric PDE problems with artificial neural networks. European Journal of Applied Mathematics, 2021, 32(3): 421-435.

Lu, T., Li, Z., Lai, F., et al. The effect of flow resistance on water saturation profile for transient two-phase flow in fractal porous media. Advances in Geo-Energy Research, 2018, 2(1): 63-71.

Lu, L., Meng, X., Mao, Z., et al. DeepXDE: A deep learning library for solving differential equations. SIAM Review, 2021, 63(1): 208-228.

Maulik, R., San, O., Rasheed, A., et al. Data-driven deconvolution for large eddy simulation of Kraichnan turbulence. Physics of Fluids, 2018, 30(12): 125109.

Meng, X., Li, Z., Zhang, D., et al. PPINN: Parareal physics-informed neural network for time-dependent PDEs. Computer Methods in Applied Mechanics and Engineering, 2020, 370: 113250.

Raissi, M., Perdikaris, P., Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 2019, 378: 686-707.

Ranade, R., Hill, C., Pathak, J. DiscretizationNet: A machine-learning based solver for Navier-Stokes equations using finite volume discretization. Computer Methods in Applied Mechanics and Engineering, 2021, 378: 113722.

Rodriguez-Torrado, R., Ruiz, P., Cueto-Felgueroso, L., et al. Physics-informed attention-based neural network for hyperbolic partial differential equations: Application to the Buckley-Leverett problem. Scientific Reports, 2022, 12(1): 7557.

Vaswani, A., Shazeer, N., Parmar, N., et al. Attention is all you need. Paper Presented at the 31st Conference on Neural Information Processing Systems, Long Beach, CA, 4-9 December, 2017.

Vinuesa, R., Brunton, S. L. Enhancing computational fluid dynamics with machine learning. Nature Computational Science, 2022, 2(6): 358-366.

Wang, S., Teng, Y., Perdikaris, P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. SIAM Journal on Scientific Computing, 2021, 43(5): A3055-A3081.

Wu, J., Xiao, H., Paterson, E. Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. Physical Review Fluids, 2018, 3(7): 074602.

Xu, R., Zhang, D., Rong, M., et al. Weak form theory-guided neural network (TgNN-wf) for deep learning of subsurface single-and two-phase flow. Journal of Computational Physics, 2021, 436: 110318.

Yang, Y., Perdikaris, P. Adversarial uncertainty quantification in physics-informed neural networks. Journal of Computational Physics, 2019, 394: 136-152.

Yang, Y., Zhou, Y., Blunt, M. J., et al. Advances in multiscale numerical and experimental approaches for multiphysics problems in porous media. Advances in Geo-Energy Research, 2021, 5(3): 233-238.