ABSTRACT
for the dissertation of Yerzhan Kenzhebek
submitted for the degree of Doctor of Philosophy (PhD) in the educational program 8D06103 – Computer Science
on the topic: "Physics-informed machine learning for modeling oil displacement problem"

Relevance of the research topic. Accurate modeling of multiphase flow in porous media is critically important for engineering applications such as petroleum reservoir simulation, groundwater management, and geological carbon storage. These processes belong to a key class of differential equations governed by the laws of conservation of mass, momentum, and energy. In particular, such conservation equations are essential for modeling filtration and transport processes, such as water- or gas-driven oil displacement in subsurface porous formations.

A classical example of such flow is water-oil displacement described by the Buckley–Leverett (BL) model. However, the equation for water saturation in the BL model has a nonlinear hyperbolic nature, which poses significant numerical challenges. These nonlinear transport problems are difficult for traditional numerical methods, mainly due to the presence of sharp saturation fronts and mixed waves (shock and rarefaction waves).

Numerical modeling methods are widely used in petroleum reservoir studies. However, in the context of limited geological data, the problem of making accurate predictions with limited information using physical laws becomes highly relevant. Traditional numerical methods (finite difference, finite volume, and finite element methods) suffer from numerical diffusion, require fine discretization to achieve high accuracy, and result in high computational costs. These limitations have stimulated the exploration of more efficient and accurate alternatives, such as physics-based machine learning, particularly Physics-Informed Neural Networks (PINNs).

Physics-based machine learning methods aim to address this problem by integrating the governing physical differential equations into the structure of neural networks. This is typically achieved by enriching the loss function with the corresponding differential equations. Currently, there is active research into exploring the potential of machine learning in modeling large-scale subsurface flow processes.

In recent years, Kolmogorov–Arnold Networks (KAN) have demonstrated superior parameter efficiency and interpretability compared to traditional multilayer perceptrons. Physics-Informed KAN models (PIKAN) show great promise in accurately representing physical laws while reducing the number of parameters and data requirements.

However, most current studies are limited to predicting water saturation within the one-dimensional Buckley–Leverett model described only by a fractional flow function. Within the PINN framework, a significant gap remains in accurately and efficiently modeling the coupled saturation and pressure fields from a unified system perspective. This dissertation proposes a system capable of simultaneously

predicting both saturation and pressure using unified PIKAN architecture, offering new momentum for modeling two-phase flow.

Objective of the dissertation

To solve the oil displacement problem based on the Buckley–Leverett model using physics-informed neural network methods (PINN and PIKAN), and to compare their effectiveness.

Research tasks:

- To review the application of PINN and PIKAN methods in petroleum production modeling
- To apply the PINN method to the oil displacement problem based on the Buckley–Leverett model
- To apply the PIKAN method to the oil displacement problem based on the Buckley–Leverett model
- To compare and analyze the results of the methods.

Research methods: the research methods include literature review, mathematical modeling, numerical methods, development and training of physics-informed neural networks (PINN and PIKAN), numerical experiments, as well as accuracy evaluation and validation of the results using statistical metrics and comparison with benchmark solutions.

Object of the research

Two-phase (water–oil) flow in porous media and the system of Buckley–Leverett equations describing them.

Subject of the research

Methods for simultaneous modeling of water saturation and pressure fields in two-phase flow using PINN and PIKAN neural networks.

Scientific novelty of the results:

- A PINN-based method was applied for coupled modeling of pressure and saturation in two-phase flow based on the Buckley–Leverett model. The combination of artificial diffusion and weighted loss functions resulted in higher modeling accuracy.
- The PIKAN architecture was adapted for coupled modeling of pressure and saturation in two-phase flow based on the Buckley–Leverett model, demonstrating effectiveness even with a significantly reduced number of parameters.

Provisions to be defended:

- Applying the PINN method to model two-phase flow based on the Buckley–Leverett model, with a combination of artificial diffusion and loss function weighting strategy, allowed reducing the mean relative L2 error of saturation by 27.86%, and of pressure by 44.4%, compared to using only diffusion.
- The PIKAN architecture enables modeling of two-phase flow based on the Buckley–Leverett model with a reduction in the number of parameters by up to 15 times (from 25,752 to 1,722) compared to the PINN method.

Theoretical and practical significance of the research

Theoretical significance: the research contributes to the development of the theoretical foundations of physics-informed machine learning methods and expands

the mathematical capabilities of Kolmogorov–Arnold networks in multiphase flow modeling.

Practical significance: the developed methods reduce computational costs in petroleum reservoir modeling and improve the accuracy and efficiency of reservoir simulation software.

Reliability of scientific results

The reliability of the results is ensured by comparison with analytical solutions, validation using traditional numerical methods, consistency with physical laws, and application of statistical analysis techniques.

Relevance to research programs

This dissertation was carried out within the framework of the following research projects:

- "Using machine learning methods to solve the problem of predicting oil production" – State registration No. 0124RK00223, project "Zhas Galym," Al-Farabi Kazakh National University, 2024–2026;

- "Parallel physics-informed neural networks for solving oil displacement problems" – State registration No. 0124RK00389, grant funding, Al-Farabi Kazakh National University, 2024–2026.

Volume and structure of the dissertation

The dissertation consists of an introduction, three main chapters, a conclusion, and a list of references. The total volume is 100 pages, including 25 tables, 39 figures, and 32 equations. The list of references contains 101 titles.

The first chapter provides a literature review on physics-informed neural networks and Kolmogorov–Arnold networks.
The second chapter presents a comparative study of PINN and PIKAN methods for predicting water saturation based on the Buckley–Leverett model.
The third chapter discusses the development and comparison of unified PINN and PIKAN architectures for simultaneous modeling of water saturation and pressure.

Publications based on the research results

A total of 10 publications were produced on the topic of the dissertation, including 3 in journals recommended by the Science and Higher Education Quality Assurance Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, 3 in peer-reviewed journals indexed in Web of Science and/or Scopus, and 4 in international conference proceedings.

- Daribayev B., Akhmed-Zaki D., Imankulov T., Nurakhov Y., Kenzhebek Y. Using machine learning methods for oil recovery prediction // Proceedings of the 17th European Conference on the Mathematics of Oil Recovery (ECMOR XVII). – 2020. – Vol. 2020. – P. 1–13. – DOI: 10.3997/2214-4609.202035233

- Imankulov T., Kenzhebek Y., Makhmut E., Akhmed-Zaki D. Using machine learning algorithms to solve polymer flooding problem // Proceedings of ECMOR 2022. – 2022. – Vol. 2022. – P. 1–9. – DOI: 10.3997/2214-4609.202244056

- Bekele S.D., Kenzhebek Y., Imankulov T. On the effectiveness of Kolmogorov-Arnold networks for enhanced oil recovery prediction in polymer flooding // Proc. of 2024 7th International Conference on Algorithms, Computing

and Artificial Intelligence (ACAI), Guangzhou, China. – 2024. – P. 1–6. – DOI: 10.1109/ACAI63924.2024.10899678

- Kenzhebek Y., Bekele S.D., Imankulov T. Prediction of two-phase flow in porous media using physics-informed neural networks // Proc. Eurasian International Scientific Conference “Artificial Intelligence and Inverse Problems in Science, Technology and Industry”. – Astana, Kazakhstan, April 14–16, 2025. – P. 215–216

- Kenzhebek Y., Imankulov T.S., Akhmed-Zaki D.Zh. Predictions of oil production by the polynomial regression method // Bulletin of KazNRTU. – 2020. – № 5(141). – P. 358–364 (in Russian)

- Kenzhebek Y., Imankulov T., Akhmed-Zaki D. Prediction of oil production using physics-informed neural networks // Bulletin of KazNPU "Physical and Mathematical Sciences". – 2021. – № 4(76). – P. 45–50. – DOI: 10.51889/2021-4.1728-7901.06 (in Russian)

- Bekele S.D., Aibagarov S., Kenzhebek Y., Imankulov T.S., Matkerim B. Predicting oil recovery factor using dense and cascade neural networks // Bulletin of the National Engineering Academy of the Republic of Kazakhstan. – 2024. – № 4(94). – P. 42–53. – DOI: 10.47533/2024.1606-146X.003 (in Russian)

- Kenzhebek Y., Imankulov T., Akhmed-Zaki D., Daribayev B. Implementation of regression algorithms for oil recovery prediction // Eastern-European Journal of Enterprise Technologies. – 2022. – Vol. 2, No. 2(116). – P. 69–75. – DOI: 10.15587/1729-4061.2022.253886 (Scopus процентиль – 54)

- Imankulov T., Kenzhebek Y., Bekele S.D., Makhmut E. Enhancing oil recovery predictions by leveraging polymer flooding simulations and machine learning models on a large-scale synthetic dataset // Energies. – 2024. – Vol. 17, No. 14:3397. – DOI: 10.3390/en17143397 (Scopus процентиль – 90, Web of Science – Q3)

- Kenzhebek Y., Imankulov T., Bekele S.D., Panfilova I., Akhmed-Zaki D. Coupled pressure and saturation prediction for two-phase flow in porous media using physics-informed neural networks (PINNs) // Computer Methods in Applied Mechanics and Engineering. – 2025. – Vol. 446, Part A. – DOI: 10.1016/j.cma.2025.118229. (Scopus процентиль – 98, Web of Science – Q1)

In addition, based on the results of the dissertation research, the following copyright object was registered: a copyright object entitled "Physics-informed neural networks (PINN) for oil production prediction" was entered into the State Register of Rights to Copyright-Protected Objects of the Republic of Kazakhstan. Certificate No. 49665 was issued on September 16, 2024. (in Russian)