Testing Gas-Liquid Homogenization for High GVF Flow Applications in ESPs

Dr. Chidirim E. Ejim, Dr. Jinjiang Xiao, Dr. Hassan M. Badr, Dr. Abdelsalam M. Alsarkhi and Dr. Rached Ben-Mansour

Abstract /

High gas volume fractions (GVFs) decrease the pressure boosting capacity of electric submersible pumps (ESPs). To prevent this, advanced gas handlers, helico-axial pumps, etc., may be installed upstream of the pump, but this equipment can be expensive. This study presents the results of testing different impeller combinations up to 90% intake GVF. The findings present a potential economical option for managing high GVF flows in ESP operation, beneficial to boosting and maximizing production from a field asset.

The pump used was a two-stage, radial-type centrifugal pump with a 3.78" impeller diameter operating at 3,400 revolutions per minute (rpm). Three impeller pairs were used: P0 (no hole in any impeller blades), Pl (holes only in the second stage impeller blades), and P2 (holes in the first and second stage impeller blades). Water flow rates were fixed from 75 barrels per day (bpd) to 550 bpd, and air flow rates varied to give intake GVFs between 10% and 92% for average fluid temperatures of approximately 25 °C. The corresponding differential pressures across the pump were measured and compared to one another.

The results showed that for all impellers, the differential pressure across the pumps decreased with an increasing GVF. At 75 bpd, pump P0 attained zero differential pressure at approximately 72% GVF. The impellers in Pl and P2 were able to extend its operation to reach zero differential pressure at 90% and 90% GVF, respectively. When the liquid flow rate was increased to 275 bpd, the differential pressures in P0, Pl, and P2 reached zero at approximately 36%, 38%, and 41%, respectively. Increasing the liquid flow rate even further to 410 bpd, results in zero differential pressure at approximately 26%, 30%, and 29% GVF for P0, Pl, and P2, respectively.

The general trend is that the GVF at which the differential pressure reaches zero decreases with an increasing liquid volume flow rate. At lower liquid volume flow rates, holes drilled in the impeller blades significantly extend the pump's GVF handling capability by homogenizing the flow at the inlet of the centrifugal pump. Since the gas handling performance of a radial-type pump was enhanced, it may be concluded that the performance will be even more favorable for a mixed flow or axial flow pump, especially at higher rotational speeds and intake pressures than in these tests.

This study highlights the importance of pursuing economical alternatives to extend the performance envelope of a centrifugal pump operating in high GVF flows. The findings from this work imply that with appropriate modifications to ESP impellers, their operating envelopes may be increased using cost-effective methods. This opens opportunities for stakeholders to maximize production from field assets with very high gas content, and increase the economic bottom line for the operator.

Introduction

Electric submersible pumps (ESPs) are artificial lift systems used in the oil and gas industry to produce fluids from weak/dead wells or increase production from naturally producing wells. The ESP pumping section consists of multiple stages, each made up of a rotating rotor (or impeller) and a stationary diffuser. Kinetic energy is imparted to the fluid by the rotors, and the kinetic energy is later converted to pressure head by the diffusers. For liquid pumping operations, the sum of the pressure rise of each stage is the total pressure rise or pressure boost by the ESP.

ESPs operate satisfactorily when pumping only liquids or liquids with very low amounts of dissolved gas. When local pressures in a well drop below the bubble point pressure, these dissolved gases come out of solution and exist as free gas with a given gas volume fraction (GVF). The GVF may be defined as the volume of free gas to the total volume of free gas and liquid at a given pressure and temperature. For conventional ESPs, the pressure boosting capacity degrades rapidly beyond a certain intake GVF value resulting in lower fluid production. In very severe scenarios, the ESP is unable to lift the fluids to the surface, resulting in gas lock[1].

The susceptibility to gas lock also depends on whether the pump geometry is radial flow, mixed flow, or axial flow. Each pump geometry is classified by a specific speed, which for radial flow, mixed flow, and axial flow pumps

is: 360 to 5,160, 1,800 to 8,260, and 8,260 to 20,640, respectively[2]. Radial flow pumps tend to have the highest susceptibility to gas lock, whereas axial flow pumps have the least gas lock susceptibility. This is mainly due to the smaller flow area, as well as the fluid inlet/outlet directions in radial pumps compared to axial flow pumps.

To delay the onset of gas lock and increase the operating envelope of ESPs pumping gas-liquid mixtures, gas handling devices such as gas separators, gas handlers, or multiphase pumps may be installed just upstream of the pumping section. Gas separators remove gas from the gas-liquid mixture, whereas gas handlers and multiphase pumps handle the gas-liquid mixture by facilitating accumulated gas entrainment and mixing to delay gas lock in the ESP. Although ESPs alone can handle intake GVFs up to 20%, gas handlers and multiphase pumps can increase the intake GVF handling up to 45% and 75%, respectively[3].

Although the mentioned gas handling devices are used in field applications, alternative methods of mixing are being postulated to facilitate gas handling in ESPs. Floyd (2008)[4] disclosed a homogenization device with a central hub and numerous posts extending from the central hub. The system may be used within gas separators or pumping assemblies. Another homogenization method was disclosed by Gahlot et al. (2017)[5]. The system consists of a diffuser with ports as well as passages through the impeller (or rotor) hub. These ports and passages allow for recirculation of the pumped fluids.

More recently, Badr et al. (2021)[6] disclosed a potential method of homogenizing the fluid in ESPs handling gas-liquid mixtures. It involves drilling holes and/or slots (both inline and in staggered configurations) through the vanes (or blades) of the impellers (or rotors). The rationale is to enhance mixing within the rotor channels from the pressure side to the suction side. The current work shows the results from testing one of the configurations in their disclosure.

This study presents the physical tests of a radial flow pump geometry operating at about 3,400 rpm. The goal was to modify the architecture of the baseline pump, and then test these modified prototypes to determine the magnitude of intake GVFs at which they will gas lock. Such information provides design knowledge, which can be incorporated in a field prototype pump as a homogenizer to extend the gas-liquid operating envelope of an ESP system and increase the economic bottom line of a field asset for the operator.

Prototypes Tested

The prototypes tested were of the radial flow pump design with two stages each, with impeller diameters of 3.78". Each impeller had four vanes (or blades), and each vane had an arc length and width of 2.362" and 0.1574", respectively. A commercial water pump with these features, and which operates at rotational speeds up to 3,400 rpm, was taken as the baseline pump, and designated P0 in this study. Modifications were made to two other identical commercial water pumps to incorporate homogenization features and then ascertain which prototype has better gas handling capabilities.

Fig. 1 *Photographs of the impellers in the two homogenizer prototypes tested.*

In the first homogenizer prototype, the first stage had no holes in its impeller vanes. For the second stage impeller, seven 0.079" holes were drilled in each vane. The holes were equidistant from adjacent holes, and at the middle of each impeller vane, Fig. 1a. This homogenizer prototype is denoted P1. The second homogenizer prototype had holes in both impellers, Fig. 1b. This homogenizer prototype is denoted P2.

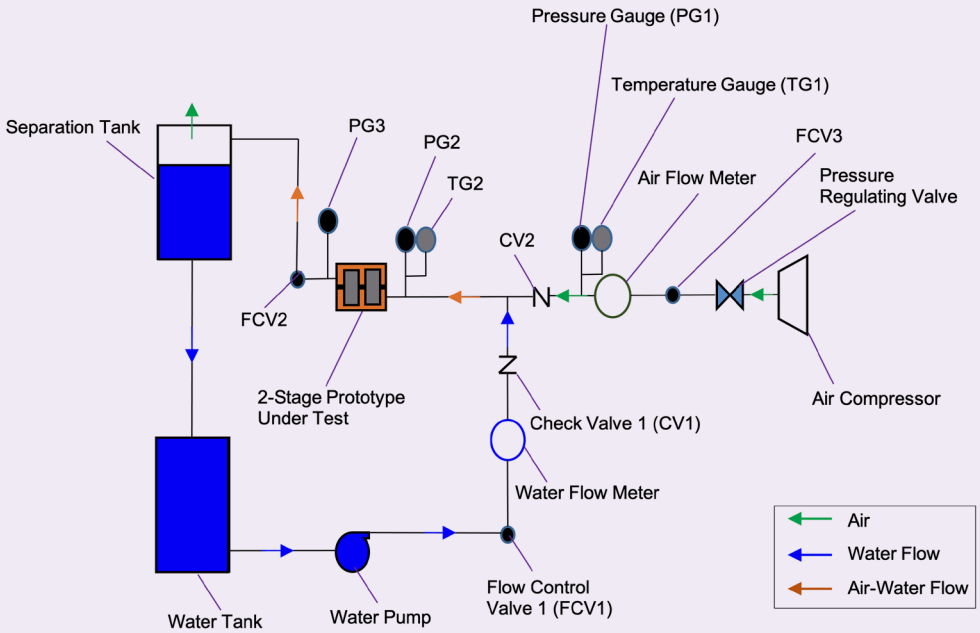
To quantify the degree of the impeller vane perforation for each of the three prototypes, a vane perforation ratio was used. This is defined as the total hole area in the impeller vanes in each prototype divided by the total impeller vane area in each prototype configuration. Based on this definition, and information provided previously on the holes and impeller vanes, the vane perforation ratio for P0, P1, and P2 were 0%, 4.6%, and 9.2%, respectively.

Test Setup and Procedure

Figure 2 is a schematic layout of the setup used to test prototypes P0, P1, and P2, with air and water as the operating fluids. Prior to initial operation of the system, flow control valves, (FCV1 and FCV2) are open, whereas FCV3 is closed. The check valves, (CV1 and CV2) prevent fluids from entering and flowing upstream of the waterline and airline, respectively. The initial operation starts by turning on the water pump and prototype (P0, P1, or P2), which causes water to be drawn from the water tank, flow through the water flow meter, and pumped into the separation tank. The separation tank is connected to the water tank via a drain line, which enables the water to flow by gravity from the separation tank to the water tank. The water in the system is allowed to recirculate for some time before air is introduced into the system.

The air supply is provided by the air compressor, whereas a pressure regulating valve allows for adjusting the air pressure into the system. With FCV3 open, air flows from the compressor through the air flow meter and mixes with the water. The air-water mixture flows into the test prototype and into the separation tank. In

Fig. 2 *A schematic of the test layout used to test prototypes P0, P1, and P2.*



the separation tank, the water and air are separated by gravity. The air is discharged back to ambient, and the water drains by gravity from the separation tank into the water tank, from where it is recirculated though the system.

Prior to taking measurements, FCV1 was adjusted to give a constant water volume flow rate, whereas FCV3 was adjusted to vary the air volume flow rates at the intake of the test prototypes. The flow conditions were allowed to stabilize before measurements were taken for each data point. The average fluid temperature during the tests was about 25 °C. For a given water volume flow rate, the corresponding air volume flow rates were measured, as well as the intake pressure (from pressure gauge (PG2)) and differential pressures (from PG3 and PG2) for the prototype being tested. The corresponding air mass flow rates and intake GVFs were also computed from this data. The air volume flow rates were increased until the pressure boost (or differential pressure) across the prototype under test was below zero, which implies the pump was in a gas lock condition.

The overall test condition can be considered isothermal given the fairly constant fluid temperature due to the relatively large mass flow rate of liquid compared to gas. The average fluid properties, list of test and measurement equipment, as well as flow rates/ranges during the test are presented in Tables 1, 2, and 3, respectively.

Results

Several observations were made following analysis of the test data. For clarity, the findings are discussed in categories related to the effects of flow parameters and the effects of the vane perforation ratio on the pump pressure boost and performance.

Effect of Flow Parameters

Figure 3 shows the variation of pump pressure boost with gas mass flow rate for P0 at liquid volume flow rates of 75 barrels per day (bpd), 140 bpd, 275 bpd, and 410 bpd. The plots show that for a given liquid volume flow rate, as the gas mass flow rate is increased, the pressure boost from P0 decreases. At a specific gas mass flow rate, the pressure boost by the pump equals zero. This scenario, which signifies gas lock conditions, is typical of pump behavior at a high amount of gas[1].

The variation in liquid volume flow rate has an effect on the gas handling capability of the pump. The figure shows that over the range of gas mass flow rates, the pressure boost by the pump at the lowest liquid volume

Table 1 *The average fluid properties during the test.*

Average Air Temperature	Average Water and Air Temperature at Prototype Inlet	Average Water Density
25 °C (77 °F)	25 °C (77 °F)	997 kg/m³

Table 2 *A list of the test and measurement equipment.*

Equipment	Range	Accuracy
Water Flow Meter	0 – 15 liters per minute	± 1.0%
	8 – 132 liters per minute	
Air Flow Meter	0 – 4 standard cubic feet per minute (scfm)	± 2.0% FS
	0 – 8 scfm	
	0 – 20 scfm	
	0 – 40 scfm	
Pressure Gauges	-15 – 15 psig	± 0.25%
	0 – 30 psig	
Temperature Gauges	-20 °C – 55 °C	± (0.04% (Reading) + 0.3 °C)

Table 3 *The test flow rates.*

Prototype	Liquid Volume Flow Rate (bpd)	Gas Mass Flow Rate Range (liters per minute/day)
P0	75, 140, 210, 275, 345, 410, 480, 550	11 to 467
P1	75, 140, 210, 275, 330, 345, 410, 480	7 to 523
P2	75, 140, 210, 275, 345, 410, 480, 550	15 to 544

flow rate tends to have a wider range of non-zero values compared to conditions at higher liquid volume flow rates. A similar trend was also observed for prototypes P1 and P2, Figs. 4 and 5, respectively.

Figures 6 to 8 show the pressure boosts from the three prototypes when plotted against intake GVF based on their corresponding intake pressures. The plots show that for a given liquid volume flow rate, the pressure boost decreases with increasing GVF. This is similar to the trend in the variation with gas mass flow rate discussed previously as well as in literature. The performance of the prototypes with increasing liquid volume flow rate is noticeable. The liquid flow rate at 75 bpd has the widest range of non-zero pressure boost values, whereas the 410

Fig. 3 *The variation of pump pressure boost with gas mass flow rate for prototype P0.*

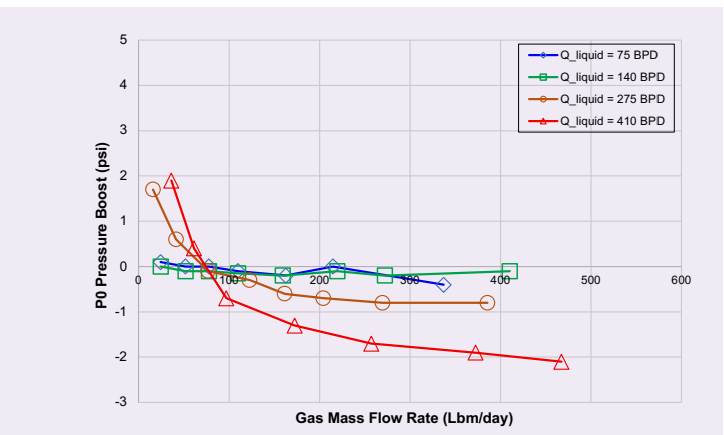


Fig. 4 *The variation of pump pressure boost with gas mass flow rate for prototype P1.*

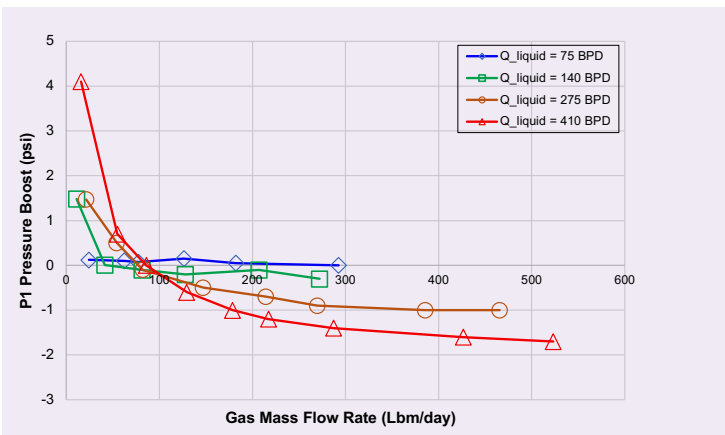


Fig. 5 *The variation of pump pressure boost with gas mass flow rate for prototype P2.*

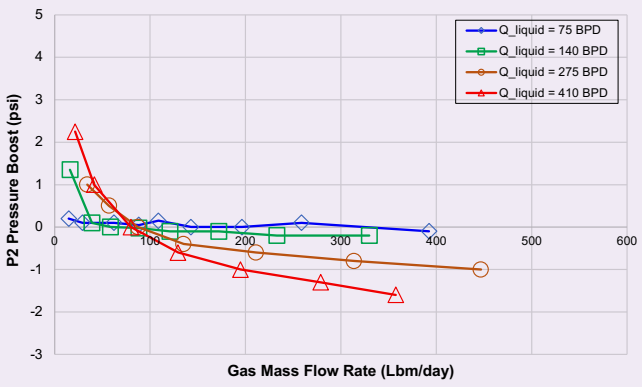


Fig. 6 *The variation of pressure boost with intake GVF for prototype P0.*

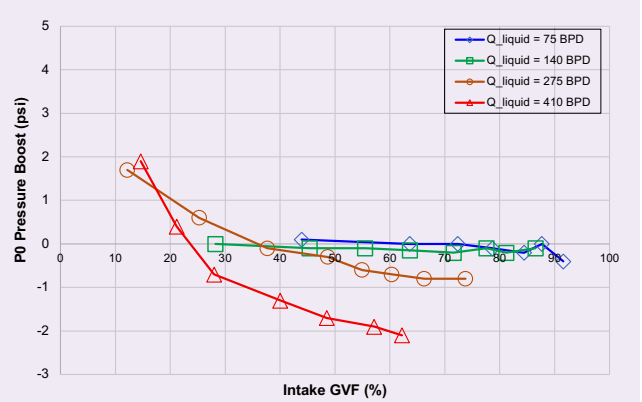


Fig. 7 *The variation of pressure boost with intake GVF for prototype P1.*

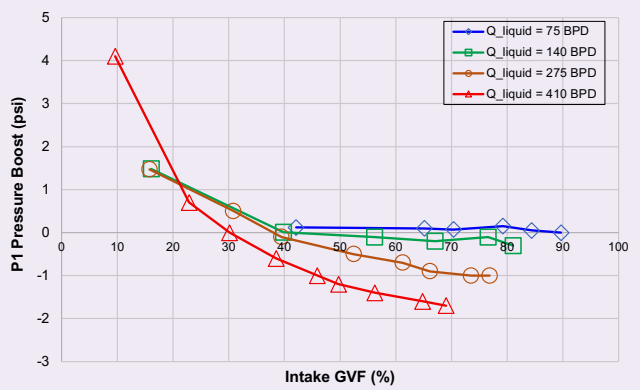


Fig. 8 *The variation of pressure boost with intake GVF for prototype P2.*

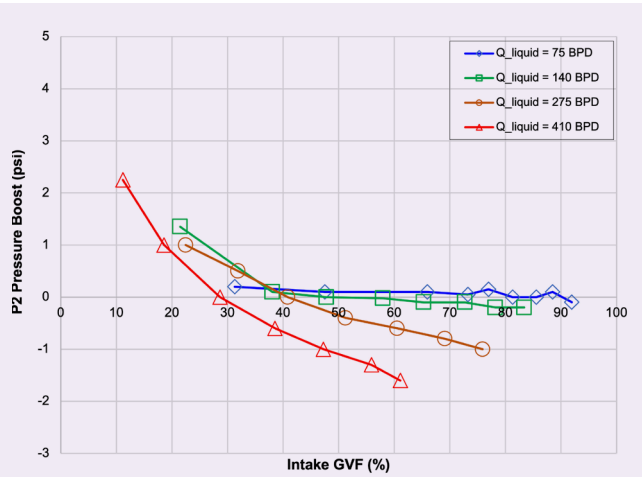


Fig. 9 *Comparing the gas mass flow rates at gas lock for the three tested prototypes.*

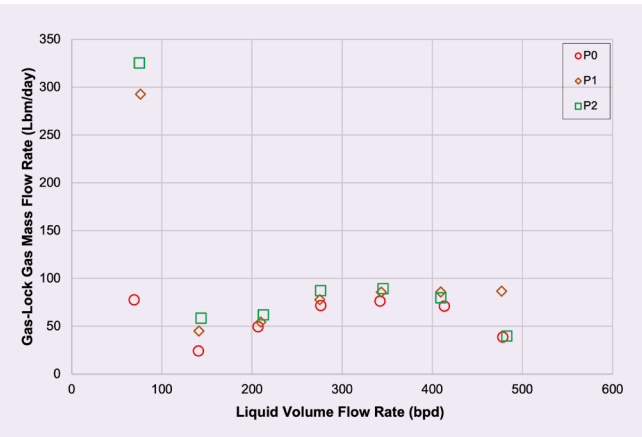
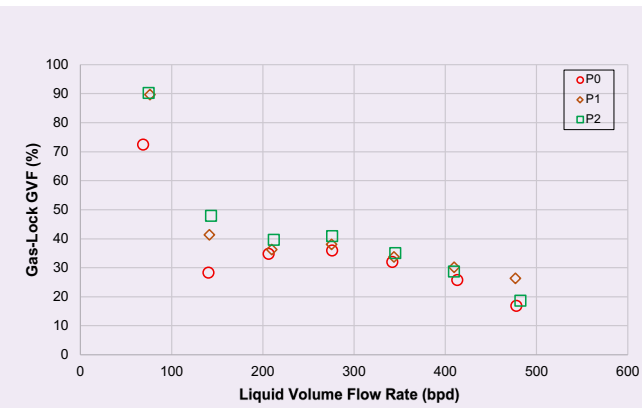


Fig. 10 *Comparing the intake GVFs at gas lock for the three tested prototypes.*



bpd data had the narrowest.

For example, for the P1 prototype, Fig. 7, when the pump was operated at 75 bpd, the intake GVF was extended up to 90% before the pump boost pressure became zero. Subsequently, when the pump was operated at a liquid flow rate of 410 bpd, the pump could only handle about 30% intake GVF before its boost pressure became zero. This behavior of attaining a higher GVF operating range at lower liquid volume flow rates may be attributed to the fundamental mechanisms of centrifugal machines. At lower liquid volume flow rates, centrifugal pumps generate higher boost pressures (or head), whereas the magnitude of the pressure boost (or head) decreases with an increasing liquid volume flow rate. The higher boost pressures (at lower liquid volume flow rates) delay the rapid decrease of the pump pressure boost to zero when pumping gas-liquid mixtures.

Effect of Vane Perforation Ratio

As previously discussed, the vane perforation ratios for P0, P1, and P2 were 0%, 4.6%, and 9.2%, respectively. The different vane perforation ratios were tested to ascertain which prototype was more tolerant to gas, and would therefore make a good homogenizer. Figures 9 and 10 show the comparisons for the three prototypes tested. The results presented are for liquid volume flow rates from 75 bpd to 480 bpd. For a given liquid volume flow rate, the data in Fig. 9 shows the corresponding gas mass flow rate when gas lock occurred for each of the three prototypes. Similarly, in Fig. 10, for a given liquid volume flow rate, the results show the corresponding GVF at which gas lock occurred.

Figure 9 shows that for liquid volume flow rates up to about 345 bpd, prototype P2 is able to handle the highest amount of gas mass flow rate before it becomes gas locked, whereas the prototype P0 handles the least amount of gas. Beyond 345 bpd, there is a negligible difference between the performance of the three prototypes, although it was observed that P1 could handle a higher amount of gas at a higher liquid flow rate. The consistently better performance of P2 for liquid flow rates below 345 bpd, may be attributed to its higher vane perforation ratio. Having more holes in the rotor vanes increases the cross-flow between the pressure and suction sides of the rotor vanes. This increases the amount of mixing of the gas-liquid mixture in P2 compared to prototypes P1 and P0.

In Fig. 10, a similar consistency in the data up to 345 bpd was also observed. Prototype P2 is able to handle the highest amounts of intake GVF compared to P1, with P0 handling the least amount of intake GVF. At 75 bpd of liquid, P2 can handle up to 90% intake GVF before it gas locks. As the liquid volume flow rate increases, its gas handling capability decreases to about 48% at 140 bpd liquid and about 35% at 345 bpd liquid. The higher amount of gas handling by P2 is due to the same reasons previously discussed for the gas mass flow rate.

Conclusions

This study was about determining which configuration of a pump with modified impellers is able to handle a higher amount of gas when pumping gas and liquid mixtures. The conclusions from the work performed include the following:

- Operating the pumps at lower liquid volume flow rates enhances the gas handling capability of the pumps by increasing the range at which non-zero boost pressure occurs. This information will aid in design and optimization of a full-scale prototype pump scaled up to meet field flow conditions.
- The prototype with the highest vane perforation ratio handled the highest amount of gas, and therefore was the best candidate as a homogenizer within the pump. Incorporating more perforations in the rotor vanes of a field prototype will therefore facilitate mixing, which extends the gas handling capability of the pump.
- The current study and tests were performed using pumps with radial flow impellers, which are not very tolerant to gas handling. In field conditions, mixed flow or axial flow pumps, which are more tolerant to gas, are typically used due to their higher flow handling capabilities. Incorporating a higher number of perforations in the vanes of these impellers will increase their gas handling capabilities even further than those observed in this study. Overall, this homogenization method offers an economic alternative to increasing the gas handling capabilities of ESPs in high intake GVF operations.

Acknowledgments

This article was prepared for presentation at the SPE Middle East Artificial Lift Conference and Exhibition, Manama, Kingdom of Bahrain, October 25-26, 2022.

References

1. Hua, G., Falcone, G., Teodoriu, C. and Morrison, G.L.: "Comparison of Multiphase Pumping Technologies for Subsea and Downhole Applications," *Oil and Gas Facilities*, Vol. 1, Issue 1, February 2012, pp. 36-46.

2. Gulich, J.F.: *Centrifugal Pumps*, 2nd edition, Springer: Berlin, 2010, 966 p.

3. Bagci, A.S., Kece, M. and Nava, J.: "Challenges of Using Electrical Submersible Pump (ESP) in High Free Gas Applications," SPE paper 131760, presented at the International Oil and Gas Conference and Exhibition in China, Beijing, China, June 8-10, 2010.

4. Floyd, R.E.: "Well Fluid Homogenization Device," U.S. Patent 7,343,967, March 2008.

5. Gahlot, V. and Tyagi, M.K.: "Multiphase Pumping System," U.S. Patent 9,624,930, April 2017.

6. Badr, H.M., Ejim, C.E., Xiao, J., AlSarkhi, A.M., et al.: "Electrical Submersible Pump with Liquid-Gas Homogenizer," U.S. Patent Application 2021/0246903.

About the Authors

Dr. Chidirim E. Ejim

Ph.D. in Mechanical Engineering, University of Alberta

Dr. Chidirim E. Ejim is a Petroleum Engineering Specialist working in the artificial lift focus area within the Production Technology Division of Saudi Aramco's Exploration and Petroleum Engineering Center – Advanced Research Center (EXPEC ARC). Prior to joining Saudi Aramco in 2014, Chidirim spent over 7 years with Schlumberger's Artificial Lift Segment, working on surface horizontal pumping systems; downhole pump design, development and testing; multiphase production systems; as well as downhole gas separator/gas handler testing and analysis.

He received his B.Eng. degree (Honors) from the University of Malta, Msida, Malta; his M.S. degree from the University of Waterloo, Ontario, Canada; and his Ph.D. degree from the University of Alberta, Edmonton, Canada, all in Mechanical Engineering.

Chidirim is a Registered Professional Engineer in Alberta, Canada.

Dr. Jinjiang Xiao

Ph.D. in Petroleum Engineering, University of Tulsa

Dr. Jinjiang Xiao is a Senior Petroleum Engineering Consultant working in Saudi Aramco's Exploration and Petroleum Engineering Center – Advanced Research Center (EXPEC ARC). His current focus is artificial lift research and development.

Prior to joining Saudi Aramco in 2003, Jinjiang spent 10 years with Amoco and later BP-Amoco, working on multiphase flow, flow assurance and deep-water production engineering.

He received both his M.S. and Ph.D. degrees in Petroleum Engineering from the University of Tulsa, Tulsa, OK.

Dr. Hassan M. Badr

Ph.D. in Mechanical Engineering, University of Western Ontario

Dr. Hassan M. Badr is an Adjunct Professor of Fluid and Thermal Sciences in the Department of Mechanical Engineering at the American University in Cairo (AUC).

He has published more than 100 papers in several refereed international journals covering a wide range of topics concerning thermofluids. Hassan has attended, presented papers, and chaired sessions in more than 40 scientific conferences. From 1990 to 2008, he was the mechanical engineering *Editor of the Arabian Journal for Science and Engineering*. Hassan is the lead author of a book titled "Pumping Machinery Theory and Practice" published by John Wiley & Sons in 2015, and a coauthor of another book titled "Oxyfuel Combustion for Clean Energy Applications," published by Springer in 2019.

He has received many awards, including the King Fahd University of Petroleum and Minerals (KFUPM) Distinguished Researcher Award (1985), Distinguished Academic Advisor Award (1988), Distinguished Instructor Award (1989), and the College of Engineering Award for Funded Research Projects (2005). Hassan has also received the Sheikh Khalifa Award for Distinguished Academic Professor at the level of Arab Countries in April 2018.

He received his Ph.D. degree in Mechanical Engineering from the University of Western Ontario, London, Ontario, Canada in 1977. Since then, he has taught many undergraduate and graduate courses in the areas of thermodynamics, fluid mechanics, heat transfer, and pumping machinery at KFUPM from 1979 to 2019, and at AUC since 2020. Hassan has also carried out several research projects in the area of thermofluids.

Dr. Abdelsalam M. Al-Sarkhi

Ph.D. in Mechanical Engineering, Oklahoma State University

Dr. Abdelsalam M. Al-Sarkhi is a Professor in the Mechanical Engineering Department and Center for Integrative Petroleum Research at King Fahd University of Petroleum and Minerals (KFUPM), where he teaches undergraduate and graduate courses in mechanical engineering.

Since joining KFUPM in 2008, Abdelsalam has been a member in several externally funded projects carried out by the Research Institute, the Mechanical Engineering Department at KFUPM, King Abdulaziz City for Science and Technology (KACST), SABIC, and Saudi Aramco.

He has published more than 100 technical papers in several refereed international journals, and a large number of reports on various areas of research related to fluid flow, multiphase flow, and thermal sciences in general. Abdelsalam received the Distinguished Researcher Award in 2015. He has attended, presented papers in several conferences related to his area. Abdelsalam remains actively involved in research projects.

He is a member of the American Society of Mechanical Engineers (ASME) and the Society of Petroleum Engineers (SPE).

In 1990, Abdelsalam received his B.S. degree, and in 1993, he received his M.S. degree, both in Thermal Sciences. Abdelsalam received his Ph.D. degree in Mechanical Engineering from Oklahoma State University, Stillwater, OK, in 1999.

Dr. Rached Ben-Mansour

Ph.D. in Mechanical Engineering, Purdue University

Dr. Rached Ben-Mansour is an Associate Professor of Mechanical Engineering in the Mechanical Engineering Department of King Fahd University of Petroleum and Minerals (KFUPM).

He has more than 25 years of teaching experience and 37 years of research experience in academia as well as industrial related projects dealing with fluid flow and thermal application issues such as flow and heat transfer in air-conditioning systems, turbomachinery, motor vehicles, industrial furnaces and boilers, thermal control of glass furnaces, simulation of fluid flow and heat transfer in petrochemical processes and manufacturing applications.

Since joining KFUPM in 2000, Rached has conducted experimental and computational research and development for vibration reduction in centrifugal pumps, computational fluid dynamic aided design of mixed flow pumps, commercial air-conditioning systems for human occupancy, as well as food storage. He has also designed and conducted full-scale experiments for vehicle aerodynamics, vehicle aeroacoustics, and air-conditioning and engine cooling systems of commercial vehicles.

Rached has successfully completed several industry and government funded projects in the U.S., Australia, and the Kingdom of Saudi Arabia. He has also collaborated with an MIT research group on developing robotic systems for gas and liquid leak detection in pipelines. Recently, Rached has been involved in the model development of hydrogen production systems.

In 1985, he received his B.S. degree, in 1987 he received his M.S. degree, and in 1993, Rached received his Ph.D. degree, all in Mechanical Engineering from Purdue University, West Lafayette, IN.

Deep Learning Model for CO_2 Leakage Detection Using Pressure Measurements

Dr. Xupeng He, Marwah M. AlSinan, Dr. Yiteng Li, Dr. Hyung T. Kwak and Dr. Hussein Hoteit

Abstract /

Geologic carbon dioxide (CO_2) sequestration (GCS) has been considered as an efficient engineering measure to decrease the global CO_2 emissions. The real-time monitoring of CO_2 leakage is an important part of big-scale GCS deployment. In this work, we introduce a deep learning-based algorithm using a hybrid neural network for detecting CO_2 leakage based on bottom-hole pressure (BHP) measurements.

The proposed workflow includes the generation of train validation samples, the coupling process of training and validating, and the model evaluation. This work solves the diffusivity equation for pressure within the Computational Modeling Group framework used to generate data sets under no leakage conditions. A Bayesian optimization process is performed to optimize the model hyperparameters. We test the performance of the hybrid neural network called the convolutional neural network (CNN) and bi-directional long short-term memory (CNN-BiLSTM) on the BHP data collected from CO_2 leakage simulations.

Results show that the CNN-BiLSTM model can successfully detect CO_2 leakage events by comparing the difference between the predicted (no leakage) and tested BHP. We further compare its superiority with CNN, LSTM, BiLSTM, and CNN-BiLSTM. Our proposed model achieves the highest accuracy with the same data sets. The CNN-BiLSTM outperforms other models owing to: (1) its capacity to process image-based input, which could accurately capture input formation, especially in cases with highly heterogeneous permeability, and (2) its bidirectional ability to capture time series dependency.

Other models, like LSTM and BiLSTM, take value-based input, which is insufficient to describe the input information in highly heterogeneous cases. In contrast, the CNN model suffers from capturing the temporal dependency features. This approach provides an efficient and practical CO_2 leakage detection method and can be implemented in large-scale GCS for the real-time monitoring applications.

Introduction

To satisfy the world's growing energy demand while reducing the effects of climate change, we should reduce greenhouse gas emissions. Carbon capture and storage (CCS) provides an attractive measure for reducing greenhouse emissions[1]. The supercritical carbon dioxide (CO_2) is injected into the subsurface during the storage process. Moreover, the presence of faults or abandoned wells may cause significant leakage risk during the CCS projects[2].

The leaked CO_2 may flow back into the atmosphere and dramatically negate the goal of the CCS project. CO_2 may also leak into other geoformations and acidize the groundwater[3]. Therefore, we should build a long-term and reduced risk CCS project. To achieve this goal, we need to track the movement of the CO_2 plume[4], and as such, various CO_2 leakage monitoring techniques have been produced in recent years[5-7].

The development of machine learning enables us to detect pressure anomalies in a data-driven manner[8]. Traditional machine learning algorithms such as the long short-term memory (LSTM) network mainly deal with time series problems and cannot incorporate spatial information. The LSTM was originally introduced by Hochreiter and Schmidhuber (1997)[9] and has been proven to be successful in many applications, such as serving as the forward model when performing history matching[10] and CO_2 leakage rate detection[11].

On the other hand, the convolutional network (CNN) was developed to deal with image-based problems. CNN is capable of extracting spatial information or features through image-based inputs and has been successfully deployed in upscaling problems[12-14]. Subsequently, CNN provides limited success in time series problems. Other neural networks such as U-Net for identifying the fractures from outcrops[15, 16], artificial neural network for gas injection optimization[17], fracture permeability estimation[18], mud loss prediction[19], the capacity of the CO_2 storage forecasting[20], and physics informed neural network and gradient-based neural network methods for solving Darcy's equations in heterogeneous media[21, 22]. The generative adversarial network can be used for digital rock reconstruction[23]. All of these mentioned applications show promising results.

Different algorithms are deployed to detect CO_2 leakage issues. Zhou et al. (2018)[24] proposed a spatial temporal CNN to detect CO_2 leakage by using seismic data. Chen et al. (2018)[25] developed a machine learning and filtering-based data assimilation approach to perform the CO_2 sequestration monitoring design. This monitoring

design process is linked with the leakage. Gundersen et al. (2018)[26] utilized the CNN to predict the CO_2 seeps behavior in water columns.

Zhong et al. (2019)[1] proposed a spatial temporal convolutional LSTM to perform the CO_2 leakage detection, where the LSTM can learn the temporal features, and the convolutional layer can learn the spatial information. De Lima et al. (2019)[27] transformed the seismic data into pseudo-red, green and blue images and utilized the transfer learning approach to detect CO_2 leakage. De Lima and Lin (2019)[28] used CNN to detect the CO_2 and brine leakage using both the pressure and seismic data. Chen et al. (2022)[29] developed reduced order models to forecast the CO_2 leakage rate, and the factors that influence the leakage rate are also investigated.

In this work, we developed a hybrid neural network called CNN bi-directional long short-term memory (BiLSTM) to detect CO_2 leakage. The CNN-BiLSTM leverages the strengths of CNN and BiLSTM, in which CNN is used for spatial feature extraction, and BiLSTM is applied for temporal dependency recognition. The BiLSTM is a type of recurrent neural network that can process the information flow in both directions. The BiLSTM can combine the present and past information by adding one more LSTM layer. For this reason, BiLSTM tends to produce more accurate predictions.

The CNN-BiLSTM enables us to build a spatial-temporal-based image-to-value regression model to learn the nonlinear mapping between high dimensional input data, e.g., permeability, porosity, injection rate, and predicted bottom-hole pressure (BHP) as an output. We deployed the Bayesian optimization process to automate the tuning of the model hyperparameters. To our knowledge, this workflow is applied in CO_2 leakage detection for the first time. The workflow provides the prediction of time dependent BHP, and the pressure anomaly is detected by comparing the difference between the predicted (no leakage) and tested BHPs.

Methods

LSTM and BiLSTM

The LSTM networks can capture short-term and long-term dependence in time series problems[9]. Figure 1a shows the structure of the LSTM. Each building unit of the LSTM includes an input gate (i), a forget gate (f), a cell candidate (g), and an output gate (o).

The BiLSTM can process the information flow from the past to the future (forward) and from the future to the past (backward), while the LSTM can only handle situations with forwarding information flow. The bi-directional ability of the information flow makes the BiLSTM different from regular recurrent neural networks. Figure 1b shows the structure of the BiLSTM, where the information can flow both in the forward and backward layers.

CNN, CNN-LSTM and CNN-BiLSTM

The CNN can take the image as input and capture the spatial dependencies of the image using different filters. The CNN can reduce the dimensionality of an image without significantly losing the critical features, which is essential for the following predictions. Moreover, CNN honors two main features: shared weights and local connectivity. The feature of local connectivity enables each neural to learn local features in a small portion, which is essential, especially when dealing with high dimensional images. The feature of the shared weights increases the learning efficiency by reducing the number of parameters. A typical CNN consists of the input, convolutional, pooling, flatten, fully connected, and output layers.

The CNN-LSTM can take the image as input and make sequence predictions. As mentioned earlier, the LSTM can only deal with time series problems while it cannot handle spatial information. On the other hand, the CNN can extract spatial information from spatial input, but fails to make time series predictions. A hybrid version of CNN and LSTM combines the strength of image-based input and time series predictions.

Moreover, CNN-BiLSTM combines the advantage of CNN and BiLSTM and therefore can process data from two directions, making it diverge from the regular CNN-LSTM. The CNN-BiLSTM has shown promising performances in various areas, such as textual sentiment analysis[30], solar radiation prediction[31], and human activity recognition[32]. Figure 2 shows the architecture of the CNN-BiLSTM.

Fig. 1 *The LSTM network architecture, where the green line represents the flow of input parameters in the unit, and the blue line denotes the flow of hidden states (outputs) at every time step. The orange line indicates the flow of cell states at every time step (a), and BiLSTM network architecture. The blue line denotes the forward flow, and the red line denotes the backward flow (b).*

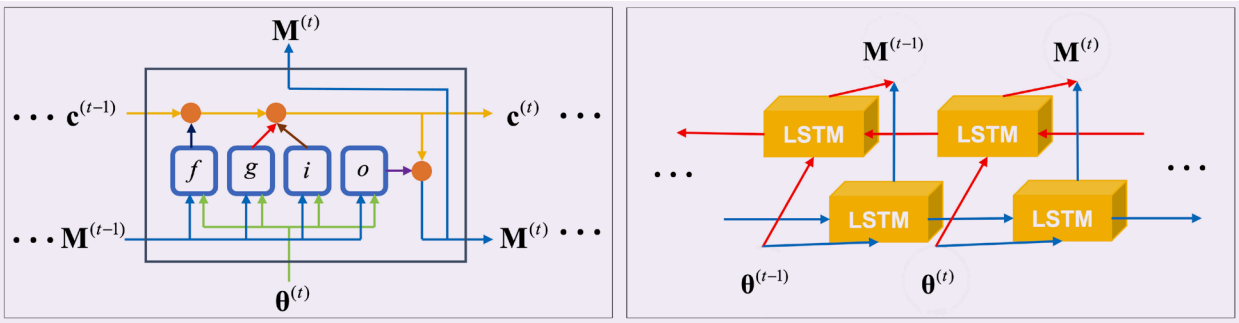
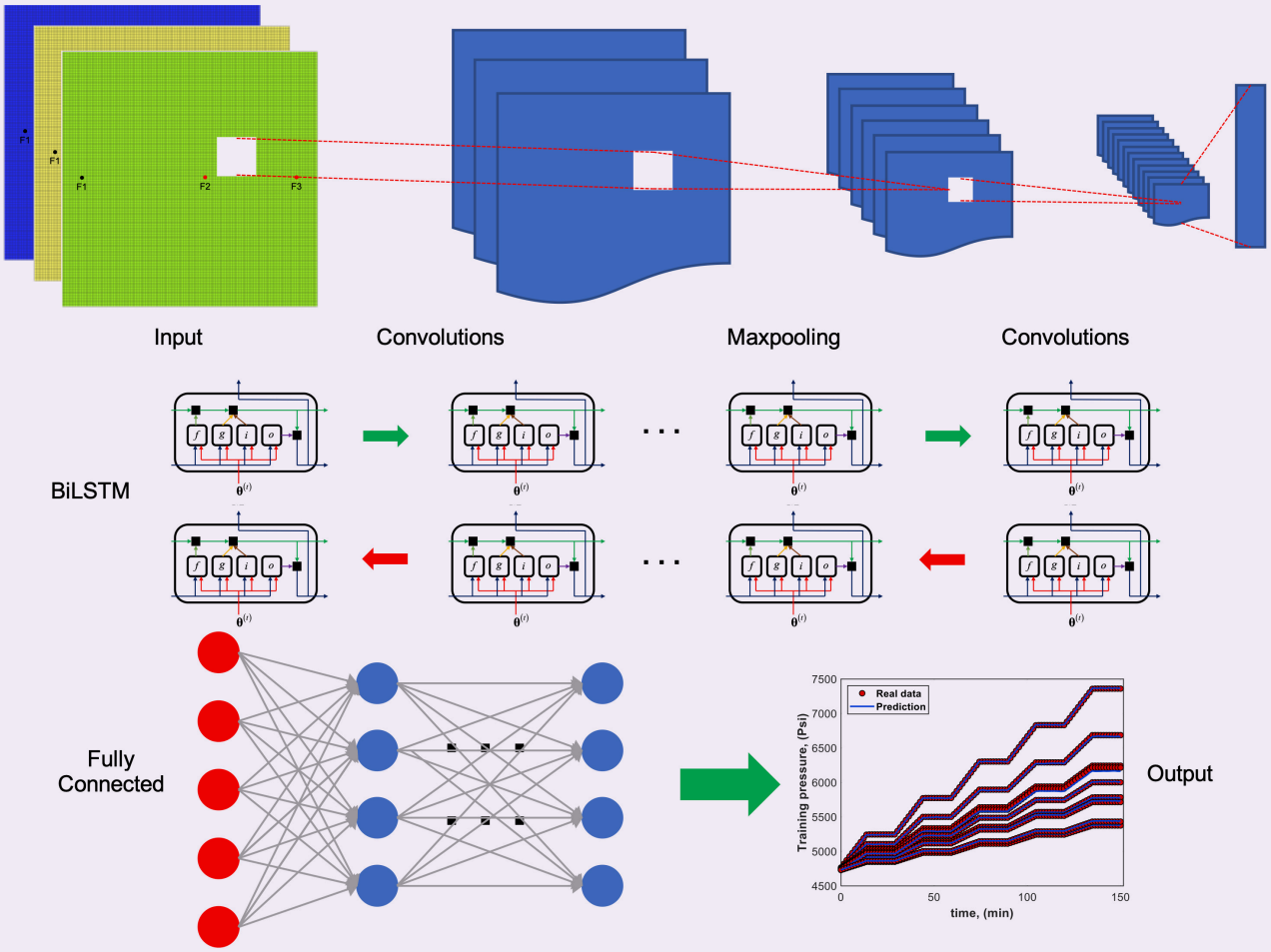


Fig. 2 *The architecture of CNN-BiLSTM showing the workflow from input data to predictions.*



Proposed Workflow

The workflow contains four main steps, Fig. 3, which are: (1) determine the uncertainty parameters and their ranges, (2) build the CNN-BiLSTM proxy model using Bayesian optimization, (3) detect the CO_2 leakage by comparing the pressure anomalies, and (4) quality check. The four main procedures are detailed here.

Step 1 – Identifying uncertainty parameters: We first identify the key uncertain parameters from existing knowledge and sample the identified parameters using Latin Hypercube Sampling (LHS). Followed by high fidelity simulation for each realization to generate corresponding output.

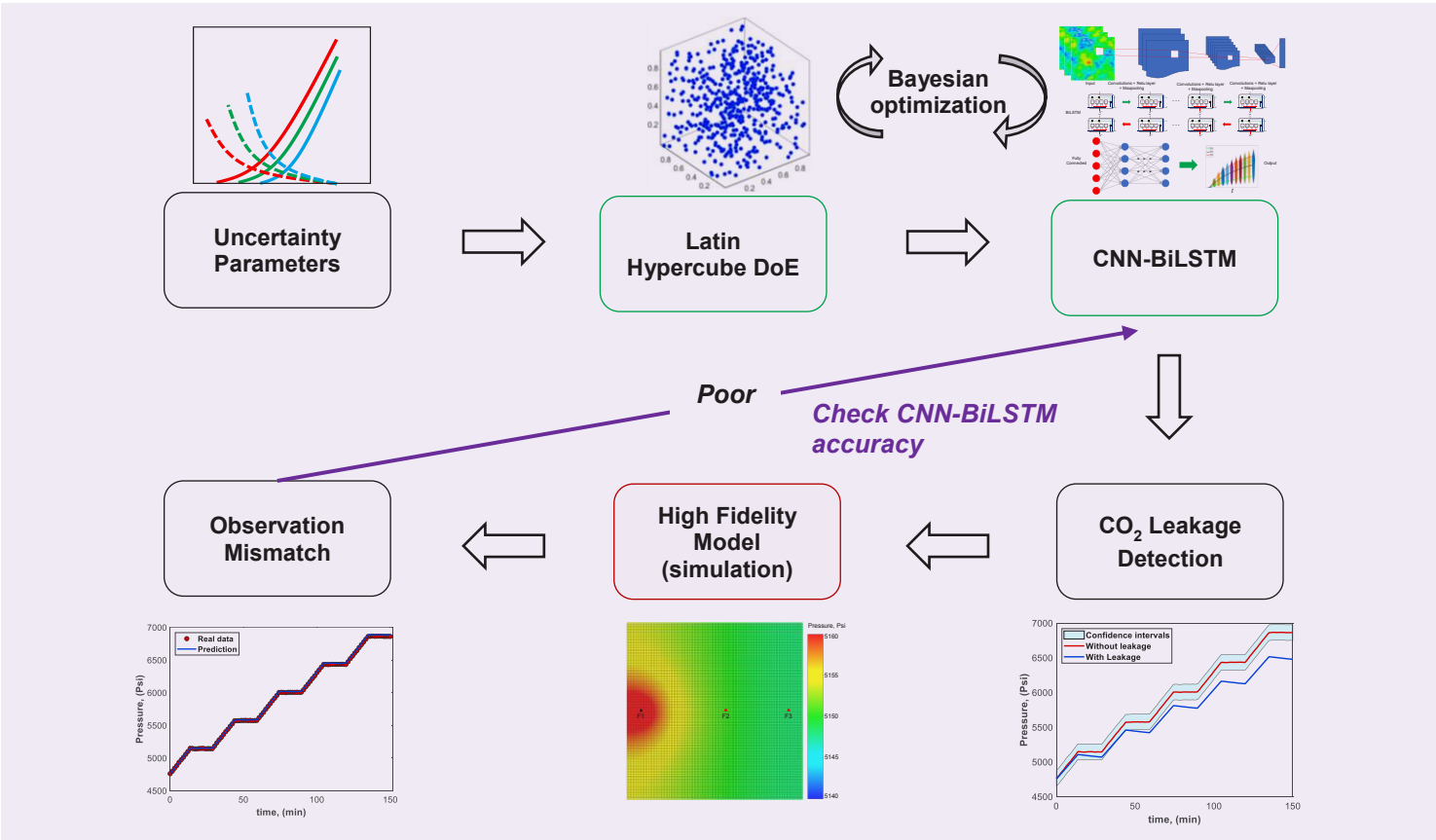
Step 2 – Building a CNN-BiLSTM proxy model: The proxy model using CNN-BiLSTM is developed to model the nonlinear mapping between the inputs and output, in which Bayesian optimization is implemented to automate the process of tuning hyperparameters.

Step 3 – Detecting CO_2 leakage: We use the well-established CNN-BiLSTM to detect CO_2 leakage. If the pressure anomalies exceed the threshold, we think the CO_2 leakage happened. To detect the pressure anomaly, we assume the pressure deviations obey the Gaussian distribution, and the detection function is expressed as:

$$p(x_i;\mu_i,\sigma_i)=\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right), \quad (1)$$

where $\mu_i=\frac{\sum_{j=1}^{N}x_i^j}{N}$, N is the length of the time series, $\sigma_i^2=\frac{\sum_{j=1}^{N}(x_i^j-\mu_i)^2}{N}$. For the new test data points, we use the detection function to determine whether the data points are within the normal range. Here we calculate the probability of the pressure difference (Δp) between the predicted pressure and the measured pressure of the monitor well. Zhong et al. (2019)[1] set the threshold to 3σ, and he pointed out that the smaller the threshold, the more reliability the trained surrogate model needs. Since we managed to reach 99% accuracy for both the training and testing data set, we set the threshold as σ in this work.

Fig. 3 An illustration of the proposed workflow.



Step 4 – Quality assessment: We then run the high fidelity model to compare with the CNN-BiLSTM predictions to check the accuracy. If the accuracy is low, we need to further improve the accuracy of the CNN-BiLSTM. This step will not be applied when dealing with field applications.

Problem Formulation

CO_2 Leakage Detection

The principle of CO_2 leakage detection using pressure measurement is that the monitor well's response diverges from the baseline when it receives a new pulse. The diffusivity equation for computing the reservoir's pressure in cylindrical coordinates can be expressed as:

$$\frac{\partial^2 P}{\partial r^2} + \frac{1}{r}\frac{\partial P}{\partial r} = \frac{\phi \mu c_p}{k}\frac{\partial P}{\partial t} \quad (2)$$

where P is the pressure, r is the radial coordinate, Φ is the porosity, μ is the dynamic viscosity of the fluid, c_p is the total compressibility, k is the permeability, and t is the time. The pressure is affected by many factors. This study considers geological formations, such as the constant porosity and permeability.

Simulation Model Description

The Cranfield area started to perform CO_2 storage in 2009. Previously, oil and gas production originated in 1944, and most production wells were abandoned in 1966[2]. The detailed area of study, where the top and bottom can be considered as impermeable layers[33], is studied. Three wells, including one injection well (denoted by F1) and two monitoring wells (denoted by F2 and F3, respectively), are located at the detailed area of study site. The effective porosity, permeability, and reservoir thickness from the previous study are used in this work[2], where the effective thickness is set to 25 m. The initial reservoir pressure is set to be 4,718 psi. The injection rate is set to 500 ft^3/d, with a cycle duration of 150 minutes, and each cycle includes a 50% shut-in period and a 50% injection period. The simulation model contains 120 × 120 × 1 grid blocks, and each with a dimension of 1 m ×

Table 1 The value and range of the base case parameters.

Parameters	Base Case	Range
Gas injection rate (ft^3/d)	500	400 ~ 600
Permeability (mD)	1.5	0.5 ~ 10
Porosity	0.1	0.05 ~ 0.3

Fig. 4 *The structure of the CNN-BiLSTM input data.*

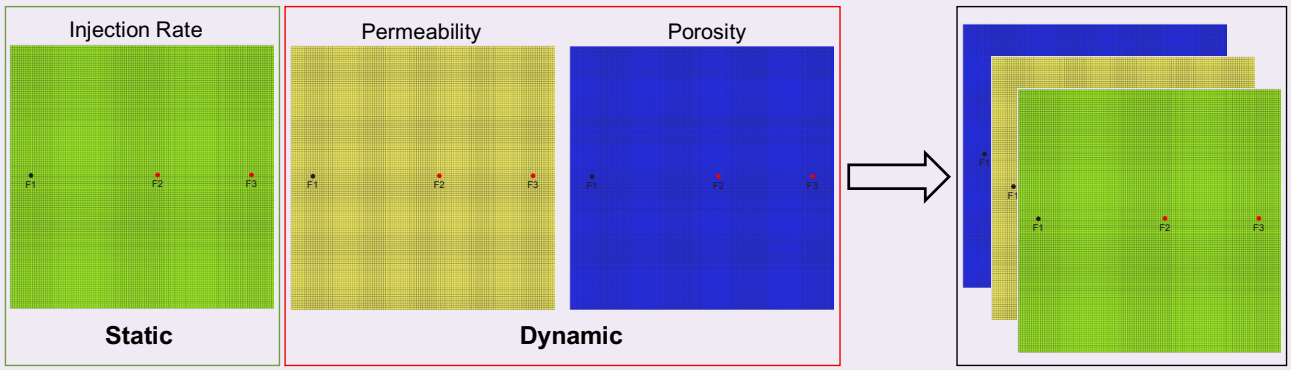


Fig. 5 *The pressure distribution after 90 minutes of CO_2 injection (a), and pressure change in Well-F2 (b).*

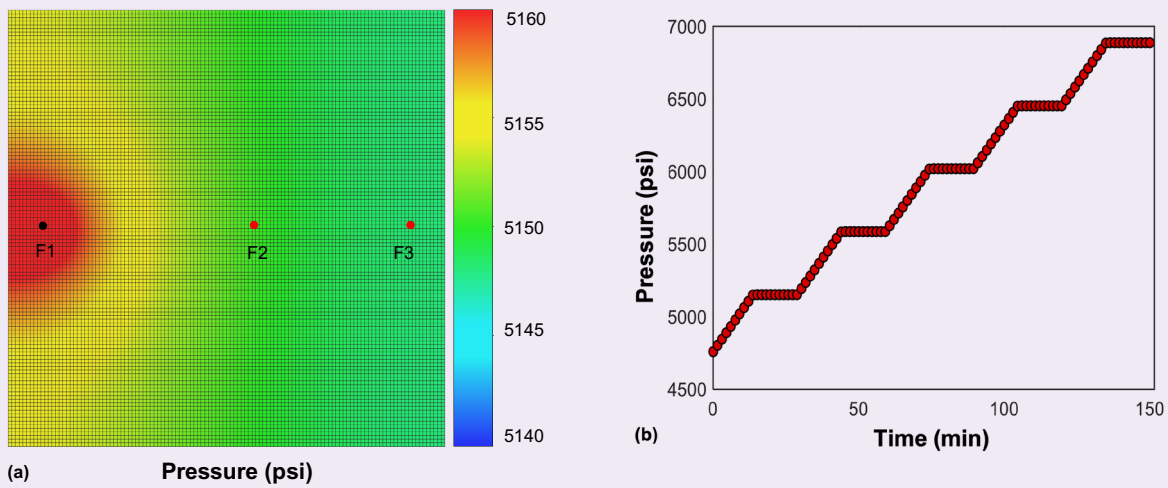


Table 2 *The initial range and the optimized hyperparameters.*

Hyperparameters	Initial Range	Optimized Hyperparameters
Neurons of the first BiLSTM layer	200 ~ 500	436
Neurons of the second BiLSTM layer	200 ~ 500	352
Neurons of the first fully connected layer	50 ~ 150	102
Epochs	100 ~ 300	251
Batch size	1 ~ 200	15
Learning rate	0.00001 ~ 0.4	0.0017145

Fig. 6 *A comparison of pressure change of Well-F2 vs. time generated by simulations and CNN-BiLSTM proxy, where (a) show cases from the training set, (b) cases from the testing set, and parity plots (c) and (d) evaluate at* t_D *= 75 minutes and* t_D *= 150 minutes, respectively*

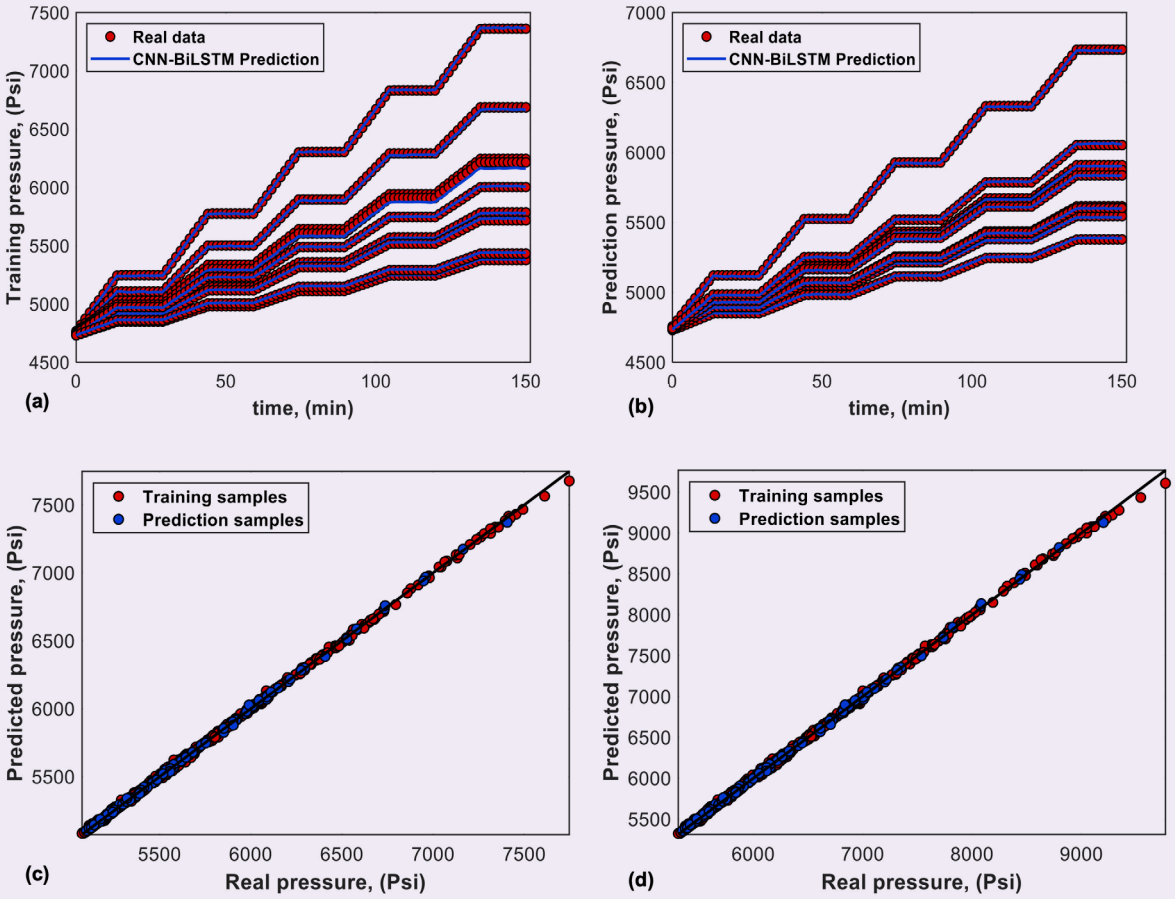
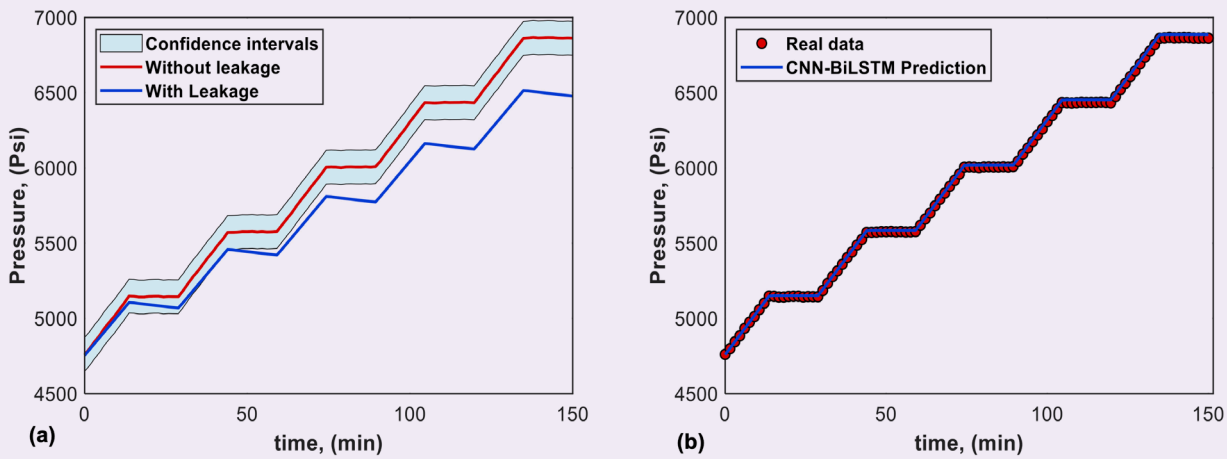


Fig. 7 *(a) The* CO_2 *leakage detection based on the CNN-BiLSTM proxy, where the red line denotes the no leakage pressure predictions based on CNN-BiLSTM, and the blue line denotes the case with* CO_2 *leakage; (b) The comparison between the pressure change predicted using CMG (red dots) and CNN-BiLSTM proxy (blue line).*



Use of Symbolic Regression for Developing Petrophysical Interpretation Models

Dr. Songhua Chen, Dr. Wei Shao, Dr. Huiwen Sheng and Dr. Hyung T. Kwak

Abstract /

A new physics guided artificial intelligence (AI) machine learning method for petrophysical interpretation model development is described. The workflow consists of the following five constituents: (1) Statistical tools, such as correlation heat maps, are employed to select the best candidate input variables for the target petrophysical equations; (2) A genetic programming-based symbolic regression (SR) approach is used to fuse multiphysics measurements data for training the petrophysical prediction equations; (3) An optional ensemble modeling procedure is applied for maximally utilizing all available training data by integrating multiple instances of prediction equations objectively, which is especially useful for a small training data set; (4) A means of obtaining conditional branching in prediction equations is enabled in SR to handle certain formation heterogeneity; and (5) A model discrimination framework is introduced to finalize the model selection based on mathematical complexity, physics complexity, and model performance.

The efficacy of the five constituent's petrophysical interpretation development process is demonstrated on a data set collected from six wells for a goal of obtaining formation resistivity factor (F) and permeability (k) equations for heterogeneous carbonate reservoirs. This study demonstrates that this new petrophysical model development process has many advantages over traditional empirical methods or other commonly used AI methods.

Introduction

For complex lithology and mineralogy rock formations, especially where the underlying correlations between logging responses and the target petrophysical attributes are nonlinear, it is difficult to come up with mechanistic or even empirical correlation equations. It is more challenging if data collected from multiple logging tools are used to build a petrophysical interpretation equation. On the other hand, most commonly used data-driven methods, such as the neural network, decision tree, K-nearest neighbor, support vector machine, or random forest, etc., deliver predicted results in numerical quantities, rather than analytical equations.

It would be extremely difficult, if not impossible, to assess whether such a solution is consistent to measurement physics with the black box prediction. Further, in the case where multiple physics measurement data are included in the input parameters, there is no transparency on how each measurement contributes to the predicted results with these black box machine learning methods.

In this article, we describe a symbolic regression (SR) approach, which is capable of combining multiphysics measurement data with genetic programming-based artificial intelligence (AI) to deliver prediction equations that utilize logging responses as inputs to capture variations of formation characteristics, thereby adjusting interpretation equations accordingly.

SR mimics genetic evolution processes of crossover and mutation to fuse different measurement physics data together for generating the target equations for various petrophysics parameters. Both measurement physics and statistical tools, such as correlation heat maps, have been employed to select the relevant, best suited input variables for a given target's petrophysical attribute.

The measurement physics is not only used in the process of selecting input variables, but also for reducing the search space of all feasible function forms, for setting constraint in fitness criteria, and for evaluating the rationality of the SR derived prediction equations. To address the limited availability of the number of core analysis data for training the models, ensemble modeling is applied with several SR predicted equations to improve the overall performance of the final prediction model.

For some heterogeneous reservoir rocks, certain petrophysical parameters often can't be described by a single equation. A traditional empirical approach is often used to create conditional branched equations with the condition that can be either facies-based or based on a key input variable. For instance, Gomaa et al. (2006)[1] proposed two permeability equations to be applied for micropore and mesopore dominated carbonates, and for macropore dominated carbonates, respectively.

Developing such conditional branched petrophysics equations empirically are quite involved and require a lot of

experience to identify the parameter used for the condition, and to quantify the condition. For SR, including or excluding conditional branch can be controlled with the inclusion of conditional operators such as max(x,y), min(x,y), etc., in the allowed pool of mathematical functions and operators for creating the targeted mathematical expressions. When they are included, the prediction equation may or may not contain the branch conditions, therefore objectively determining the necessity of the choice of conditional branching.

Black box AI predicted results often have the deficiency of being unable to compare two models predicted with different methods, different data, or different termination conditions, other than just a model performance such as R^2 of the verification data vs. the model predicted values. With the SR delivered prediction equations, rather than predicted values, one is able to evaluate two models beyond the model performance alone; one can also examine the different models based on their mathematical complexity and physical measurement complexity.

The five constituent's petrophysical interpretation development process described in this work has been tested on a data set for the objective of obtaining formation resistivity factor (F) and permeability (k) equations for a heterogeneous carbonate reservoir. Logging responses equivalent core analysis data, including nuclear magnetic resonance (NMR) T_2 distribution and its derivatives such as T_{2GM}, free fluid index (FFI), micropore and macropore volumes, ϕ_{micro} and ϕ_{macro}, total porosity ϕ, and acoustic measurements, V_p, along with the core perm and F as the "ground truth" target data, are used to train the interpretation equations.

Following the correlation heat map ranking, the highly correlated inputs are used with the SR to generate models that are better than the Archie equation for the case of F and cementation factor, m, where variation of m can be derived based on porosity, V_p and ϕ_{macro}. For the k equation, the inclusion of porosity, T_{2GM}, and ϕ_{macro} have proven to perform superior to other combinations of input variables. These demonstrate that SR derived prediction models are capable of building the variation of formation via logging responses into the interpretation models.

Methodology

Review of Conventional Empirical Approaches

When one needs to derive petrophysical interpretation models for a new, heterogeneous reservoir rock formation, petrophysicists usually take one of the following two approaches. Often, one would start from well-known equations, such as Archie's equations[2] for saturation or SDR[3] or Coates equation[4] for k, honoring the general function form and the input dependencies of the equations, but tweaking the coefficients and/or model parameters.

For instance, for Archie's formation resistivity factor (F) and saturation (S_w) models,

$$F = \frac{R_0}{R_w} = \frac{a}{\phi^m} \quad (1)$$

and

$$\sigma_t/\sigma_w = \phi^m S_w^n/A \quad (2)$$

the general power-law function forms of the equations are kept, while the cementation exponent m, saturation exponent n, and/or tortuosity parameter a are "calibrated" to fit the specific formation. In Eqns. 1 and 2, ϕ is the total porosity, R_0 is the resistivity of brine-saturated rock, and R_w is the resistivity of the brine fluid. σ_t and σ_w are the conductivity of the total fluid saturated and brine, respectively. Similarly, for the SDR[3] k model,

$$k = C(\phi_{NMR})^a (T_{2GM})^b \quad (3)$$

the parameters of $a = 4$, $b = 2$, and C are "calibrated" formation dependent parameters while the general power-law dependency function form is kept untouched. We call such an approach the "parameter tweaking approach."

Conceptually, a parameter tweaking approach acknowledges the difference of one formation rock type from another, however, one single set of "calibrated" parameters may not be adequate to deal with the local heterogeneity of rocks. When one set of parameter tweaking is insufficient, petrophysicists usually favor the facies-based parameter tweaking approach, where the local variation is identified as different facies, and for each facie, a set of parameters is obtained from a core calibration. Though an improvement, the facies identification itself usually is subjective and a discrete number of facie IDs may be still insufficient to account for depth-by-depth variation effectively.

The parameter tweaking approach has a fundamental assumption of the validity of the original equation form. For instance, Archie's equation is based on a so called Archie's rock. Complex rock formation may depart from Archie's rock type characteristics significantly, such that the validity of the Archie function form is no longer applicable. When this occurs, petrophysicists would seek for new equations, which may include different measurements and different function forms. This second route carries a lot of unknowns, and therefore, naturally, it is in a less comfortable zone for many. Not only is the optimal equation form unknown, the measurements that might be useful as input variables are also unknown. On top of that, local heterogeneities may require different equations to be used.

Objectives of New Model Development Approach

Our objective is to develop a machine learning approach that has the following features:

- Use well logging data as input to the petrophysical models to capture the continuous reservoir rock heterogeneity depth-by-depth, although core analysis data are needed to train the models.
- Predict target petrophysical equations rather than black box values.
- Be capable of predicting the target equations of either type: the "parameter tweaking" for the existing equations, i.e., predicted equation for parameter variations, or a new equation not related to the existing equation form with the same or different measurements.

- Combine multiphysics logging measurement data inputs objectively in seeking a target equation.

Workflow Description

Correlation Heat Map for Multiphysics Data Selection

The first task in any petrophysical model development process involves selecting the most relevant tool measurement data (model input variables) to be used to train the model. Tool measurement physics is usually the first to consider. Measurement physics tells the dependency of a particular tool's deliverables in general, but for heterogeneous rocks, the sensitivity of such dependency is not always obvious. Therefore, in our process, the measurement physics criteria are used primarily for a more inclusive selection of data. Next, we use a statistical method to decide which data are to be included as the input variables for a target model.

The statistical tool we use is the correlation heat map. Two types of heat maps have been routinely used in our workflow: Pearson's linear correlation and Spearman's monotonic ranking correlation. Figure la is an example of a linear correlation and Fig. lb is an example of a monotonic ranking correlation.

Figure 2 shows an example of heat maps based on the Pearson correlation, and the Spearman correlation for a target parameter of permeability, k, with several measurement variables, including sonic compression velocity, V_p, ϕ, T_{2LM}, ϕ_{micro}, ϕ_{macro}, and *FFI,* respectively. The color code represents the correlation strength from 1 (strongest positive correlation) to -1 (strongest negative correlation). A correlation coefficient in the neighborhood of zero represents the weakest correlation. Ideally, we shall prioritize the selection of input variables (measurement data) only with stronger correlations (< -0.7 or > 0.7)

Fig. 1 *An example of a linear correlation (a), and a monotonic ranking correlation (b).*

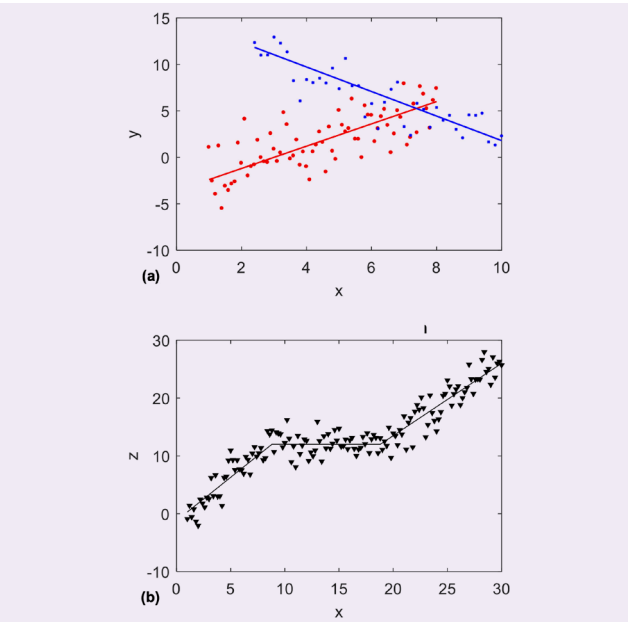
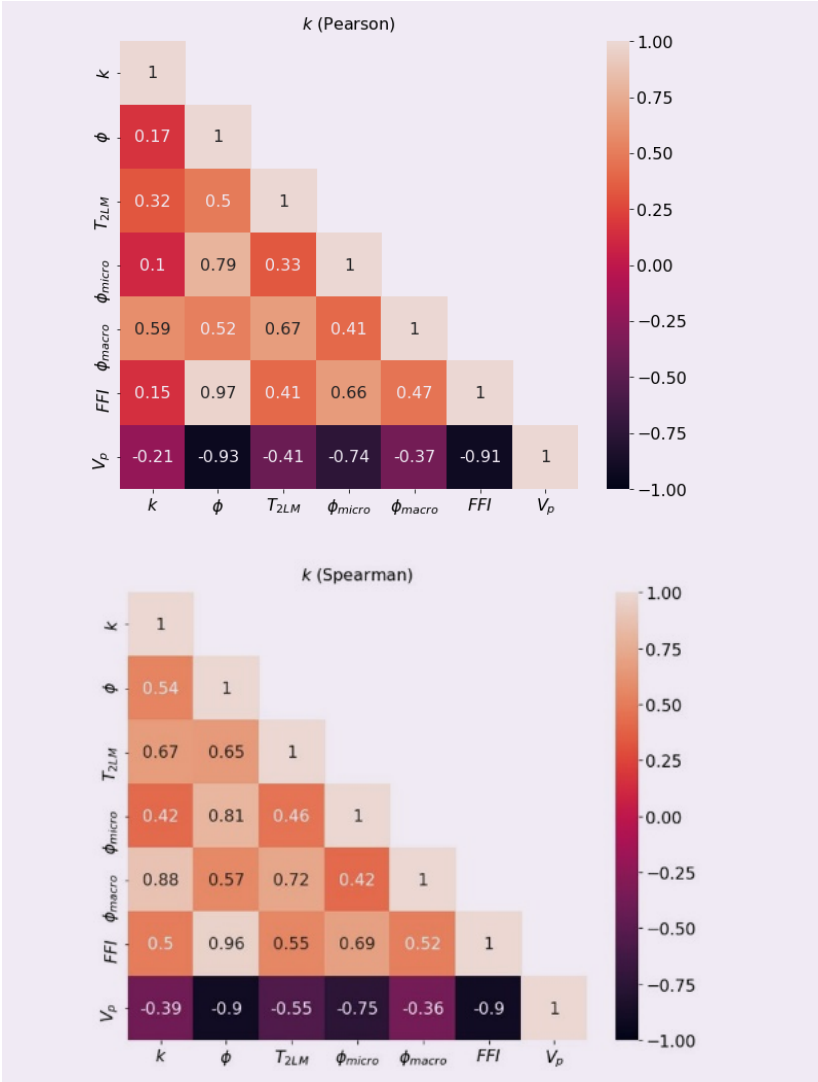


Fig. 2 *Examples of the Pearson heat map (top), and the Spearman heat map (bottom).*



but the criteria can be relaxed based on the availability of the strongest correlation data.

A high Pearson correlation often implies a high Spearman correlation (an exception exists which will be discussed later) but the vice versa is not always true. Therefore, if we found very high Pearson and Spearman correlation coefficients for one or more candidate input variable(s), this implies simple linear regression or multilinear regression may be sufficient to build a petrophysical model, even though SR still can be used. On the other hand, if the Spearman correlation coefficient is significantly higher than the Pearson correlation coefficient, this implies a more complicated model equation may work better, in which the SR method is expected to produce the prediction equation better.

SR

The SR analysis algorithm searches the mathematical expressions to come up with the prediction model

equation that best represents a given training data set, in the defined accuracy and simplicity criteria. Among all available SR software packages, DataRobot[5] and AI Feynman[6] are most popular. The work in this article is based on the use of available algorithms in DataRobot. This does not imply one is more favored than the other.

SR in DataRobot is a genetic programming based algorithm. Genetic programming resembles nature's evolution processes of crossover and mutation. Crossover pairs the genetic materials from each parent to cross over from one another to pass to the child, which can be viewed as a combining process. A mutation occurs when the genetic message carried by the gene of the parent is altered to pass to the child. In our model development process, the primitive function expressions and operators are the parent genes and the SR algorithm follows the genetic processes to create the prediction function. Figure 3 is an illustration of this process.

In this process, a multiphysics tool data as an input variable is fed into the genetic programming algorithm, along with preselected functions and mathematical operators, as well as predefined fitness objects. The preselected functions and operators narrow the "gene pool," and measurement physics knowledge would be very useful to exclude those that do not obey the measurement physics, and add the ones that are known, correlating well with the target.

For example, triangular functions are largely irrelevant to many petrophysical interpretation equations, therefore, they are better excluded to avoid the risk of fitting noise. It is also desirable to combine certain input variables in accordance to their correlation with the target variable. If $\frac{\phi_{micro}}{\phi}$ is known to correlate to k better than ϕ_{micro} and ϕ separately, it would make sense to define $\frac{\phi_{micro}}{\phi}$ as one of the preselected input variables, along with ϕ_{micro} and ϕ. Doing so increases computational efficiency and makes the target equation more physically interpretable.

It should be noted that customizing input variables (such as $\frac{\phi_{micro}}{\phi}$) do not cause the predicted equation to be subjective to user bias. The SR prediction process just treats it as one independent variable, which may or may not appear in the predicted outcome — the target equation.

Iteratively, the genetic evolution process in SR runs repeatedly until a determination condition is met, which is either when a maximum number of generations has been produced or a satisfactory fitness level has been reached. Then, the target prediction equation is obtained. It is important to understand that genetic evolution processes and crossover and mutation, occur randomly. This random nature is a key feature in the SR process. The randomness ensures that the outcome is not subject to human bias and the combination of multiphysics measurements into one equation is naturally optimized.

On the other hand, the prediction outcome is not expected to be unique, namely if we run the same data with the same predefined fitness objects multiple times. It is no surprise that we may obtain different prediction equations, as the result of a random genetic programming process occurs in each run. When we have a large, diverse training database, these different equations should have equivalent effectiveness, and they are all valid.

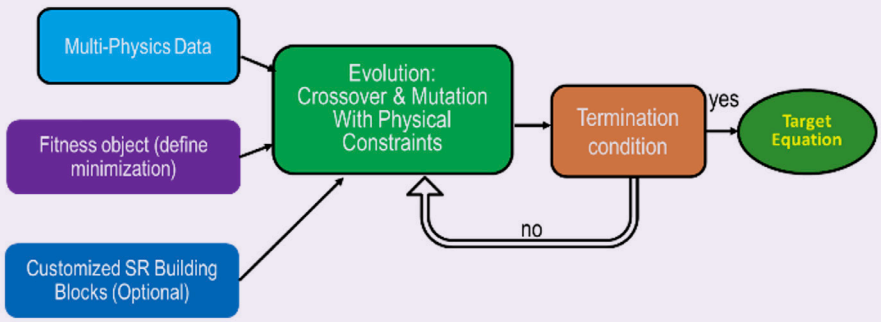
Randomness also occurs in the inclusion of training data when running SR in multiple instances independently. During the SR model training process, the DataRobot algorithm randomly selects a subset of all available data, e.g., two-thirds, for training, and the remaining one-third for validation. With each training run, the training data are randomly selected independently. Again, this ensures unbiased training, but a nonunique, though equally valid, outcome.

Ensemble Modeling

When the training data sets are large, taking two-thirds of the data set is still a large database, therefore it is unlikely to affect the outcome's general representativeness to the entire database significantly. Consequently, when the training data set size is marginally acceptable, each training run's result could defer in the performance. It is desirable to come up with a method that can combine multiple runs' diverse equations in an objective manner, thereby the resultant ensemble model aggregates the prediction of each base model used as the final prediction model.

The concept of ensemble modeling is applicable to base models that are diverse in training data, diverse in training algorithms (random genetic programming iteration can be viewed as a different training algorithm in a broad sense), or the combination of the two, which

Fig. 3 *An illustration of the SR process for predicting petrophysics equations.*



makes it applicable to SR trained base models. In general, ensemble modeling can be expressed as:

$$M_{ens} = \sum_i c_i M_i \quad (4)$$

where M_i is the *i-th* base model, c_i is the weight coefficient of the *i-th* base model, and M_{ens} is the resultant ensembled model.

The formation of an ensembled model is not limited to the way described in Eqn. 4. For instance, if all the base models are in a power law form:

$$M_i = \prod_k A_{i,k} x_k^{s_k} \quad (5)$$

it may be desirable to construct the ensemble model in logarithmic scale:

$$\ln M_{ens} = \sum_i \ln x_i \quad (6)$$

where x_i is *i-th* input variable and $x_0 = 1$ is a defined constant.

Conditional Branching

Conditional branching is a model development technique that is especially applicable to scenarios where one or more input variables correlates to the target prediction parameter in distinct clusters. In a conventional empirical model development process, usually it takes an experienced petrophysicist to discover such branching criteria, and in the process, may introduce human "preference" or bias. Most data driven prediction methods, at best, may need to first use a clustering approach to identify the clusters, then try to come up with individual equations for each cluster. This is different from the direct find optimal branching condition during the training process.

While conditional branching can be an effective means to optimize a prediction equation, the overuse of conditional branching could potentially cause overfit with the piecewise fitting mechanism, resulting in an increased uncertainty of model application envelop to new data. Therefore, this capability is only switched on when the performance of the unbranched prediction equation is not satisfactory.

The conditional branching capability in genetic programming-based SR can be turned on and off by inclusion or exclusion of branching enabled operators, such as max or min operators, providing us easy control.

When the conditional branching is enabled, the tendency is that the prediction equation is frequently to contain a branch, even though it may not be necessary, as a branching usually improves fitting of data. To prevent unnecessary conditional branching, we developed the fifth feature of our workflow to objectively determine whether the branching is necessary.

Model Selection Criteria

A branched equation increases model complexity. Therefore, it is desirable to compare several prediction model forms for its complexity. Initially, the need of such criteria arises from comparing branched and unbranched models. Soon, we recognize the same concept is useful to compare the models from various SR random runs, and even SR delivered models with other existing empirical model equations.

In consideration of a broad application model selection criteria, equation complexity alone is insufficient. Different models may contain different input variables, which may come from different logging instruments, and which have different data fidelity, availability, and cost to acquire issues. In addition, model performance, such as the R^2 error, may need to be considered. Along these considerations, this leads us to believe an effective optimal model picking process should include three elements: (1) mathematical complexity, (2) physical measurement complexity, and (3) model performance. Each can be assigned a score. The lower the overall complexity of the score, the better.

A mathematical complexity score usually can be evaluated with known algorithms. Su et al. (2021)[7] described an algorithm for the computing mathematical complexity of a mathematical expression E as the following:

$$\Gamma(E) = \alpha\Gamma_r(E) + \beta\Gamma_c(E) \quad (7)$$

where Γ is the representation complexity of E, and $\alpha\Gamma$ is the computation complexity of E. The detail of $\Gamma(E)$ and $\Gamma_c(E)$ can be referred in Su et al. (2021)[7].

We adapted a much simpler algorithm for computing the mathematical complexity of a mathematical expression E. The algorithm represents a mathematical expression E in a binary tree. The depth of the binary tree for the mathematical expression E is considered as the representation complexity of E. Each operator or function in the binary tree is assigned a score for its computation complexity. For example, + and − operators are assigned a score of 0, /, and * operators are assigned a score of 1, etc. The computation complexity of E is computed cumulatively by traversing the binary tree.

Application Examples

The five constituent's workflow previously described has been applied to a data set containing core data from five carbonate wells. All the input variables are from core analysis measurements. These core measurements are logging equivalent variables, meaning that all of these core analysis deliverables have an identical counterpart in logging measurements. Therefore, the models developed with training data from core analysis can be applied directly to log interpretation. The target petrophysical parameter in this study is F, and k.

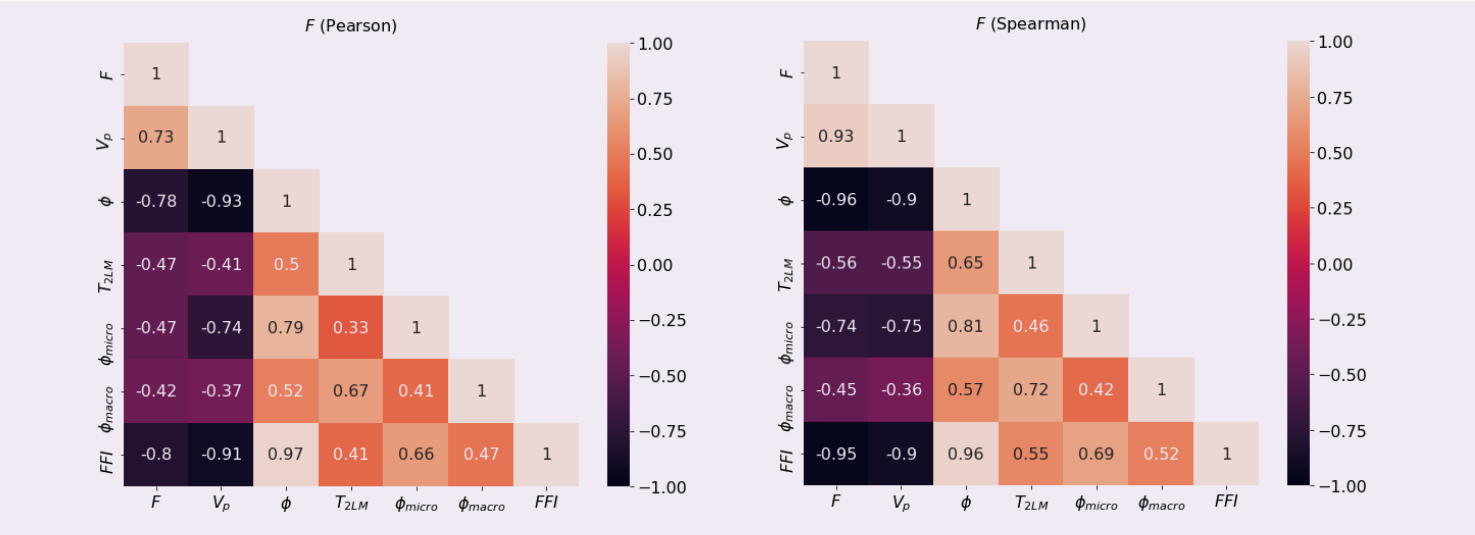
F Prediction

In this example, we demonstrate the use of a subset of the five constituents to predict F for capturing the formation heterogeneity.

Figure 4 shows the correlation heat maps for F. Both Pearson and Spearman correlation coefficients indicate that porosity, ϕ, compressional velocity, V_p, and NMR *FFI* are quite strongly correlated to F. In addition, NMR T_{2LM} is also a potential candidate, considering the Spearman ranking correlation coefficient is greater than 0.5.

For convenience, it is also preferred that a sorted heat map column is used with the target parameter on the top, followed by the input candidate variable having the

Fig. 4 *The Pearson (left) and Spearman (right) heat maps for* F *with several input variables.*



strongest positive correlation with the target. Other candidate input variables are listed in the descending order according to the value of the correlation coefficient with the strongest negative correlation on the bottom, Fig. 5.

The heat map plotted in this manner allows us to quickly spot the strongest candidate input variables (both positive and negative) to be included in the training. The sorted Pearson and Spearman heat maps, plotted side by side, also conveniently show the difference between the two. Often, the Spearman correlation is higher than the Pearson. Bucking this trend usually indicates that the candidate input variable's distribution may be farther away from normal distribution and weighs more in the tails. This may not be a negative indication, but certainly could be influential for us in designing the training, i.e., to justify the conditional branching parameter.

While the heat maps presented in the form shown in Fig. 5 have certain benefits, the original heat maps shown in Fig. 4 present additional correlations between multiple candidate input variables that the sorted heat maps cannot. If two of the candidate input variables are strongly correlated, including both may not be necessary. On the other hand, if two candidate input variables all highly correlated to the target, but they weakly correlate to each other, including both can be beneficial for developing the model because they are independent.

Several different training strategies were used and are described next. In the first simple attempt, we tried to demonstrate that SR can reproduce the Archie model. When we narrow the input variable to porosity only, the SR indeed produced a power-law dependence equation from the training date set:

$$F = \frac{a}{\phi^m} = \frac{2.53}{\phi^{1.40}} \qquad 8$$

The cross plot of the predicted vs. measured F for this carbonate field is illustrated in Fig. 6, with $a = 2.54$, $m = 1.40$, and $R^2 = 0.74$. The available data set only contains 32 samples, two-thirds are randomly allocated for training with one-third for validation. With such a small data set, each of the individual runs could produce somewhat different results, i.e., m and a in Eqn. 8 can vary somewhat.

Next, we demonstrate that the heterogeneity of the formation on F determination can be incorporated by tweaking the cementation factor. Namely, we honor the

Fig. 5 *Other candidate input variables listed in the descending order according to the value of the correlation coefficient with the strongest negative correlation on the bottom.*

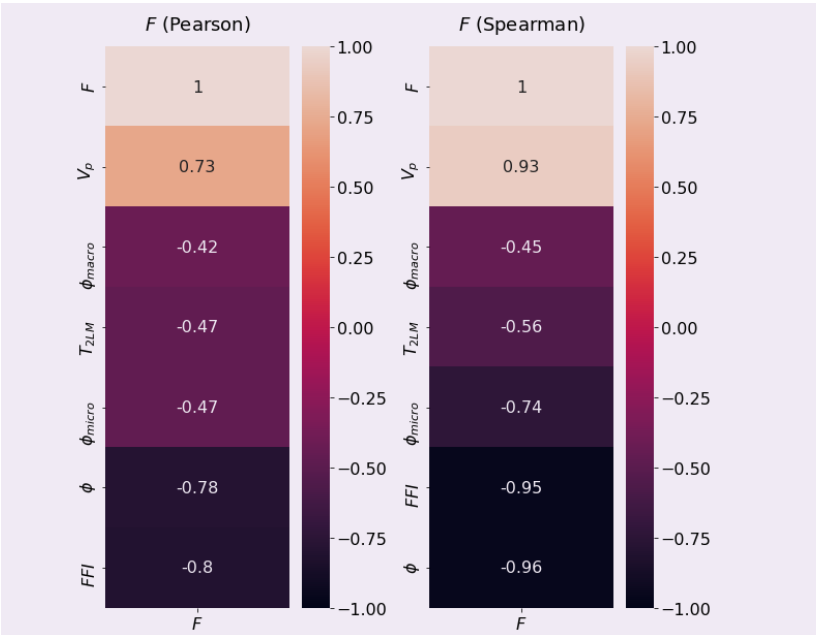
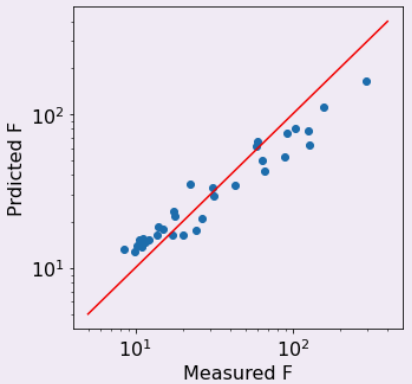


Fig. 6 *The cross plot of the predicted vs. measured* F *for this carbonate data set with* R_2 *= 0.74.*



Archie equation form but allow m to be variable and m is determinable by independent logging-data-equivalent measurements. The SR predicted result is:

$$F = \frac{2.59}{\phi^m} \quad (9)$$

where

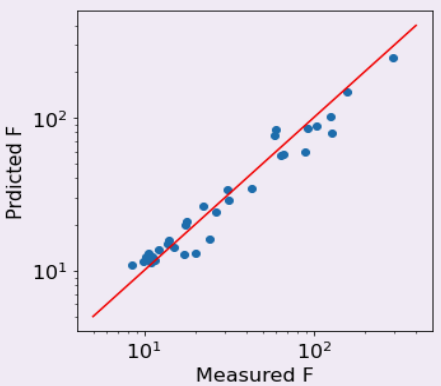
$$m = 0.17 \ln T_{2LM} - 0.29 \ln \phi. \quad (10)$$

The predicted F (by Eqn. 10) vs. measured F cross plot is shown in Fig. 7, with $R^2 = 0.93$.

Equations 9 and 10 are trained based on the randomly allocated two-thirds of the data and the other one-third for validation. We can see that by allowing m to be an equation with variables of T_{2LM} and ϕ, the model performance is much improved.

To further demonstrate that model Eqn. 9 with Eqn. 10 is adequately optimal, the ensemble modeling concept is applied and refitting data with all instances. The resultant equation is only minorly different with:

Fig. 7 *The SR developed equation for* F *by tweaking the Archie parameter, allowing the cementation exponent variable and determined by independent logging equivalent measurements,* R_2 *= 0.93.*



$$F = \frac{2.09}{\phi^m} \quad (11)$$

and

$$m = 0.2 \ln T_{2LM} - 0.29 \ln \phi. \quad (12)$$

We see the coefficients in Eqns. 11 and 12 only have small differences from those in Eqns. 9 and 10, indicating the equation is robust because it is insensitive to the random selection of training data, and $R^2 = 0.96$ also shows insignificant improvement from that of Eqn. 10.

We have demonstrated the use of SR to predict F by tweaking the Archie parameters. The tweaking of the model is much more than just finding a different constant to replace the default a and m. Using Eqn. 10 or Eqn. 12, the effect of formation heterogeneity on F is captured by T_{2GM} and ϕ, which allows on-the-fly continuous variation of the Archie model. Such continuous variation of the Archie parameters is much more powerful than facies-based discrete sets of Archie parameters. Moreover, it does not require human interference of determining facies first.

Next, we demonstrate applying SR on the same set of data to predict the F equation without being constrained by Archie's equation form. As we previously described, the randomness nature of the SR process may yield diverse prediction equations when running it multiple times. Two of such prediction equations are:

$$F = \frac{V_p^{1.38}}{e^{10.52} FFI^{0.37} \phi^{0.91}} \quad (13)$$

and

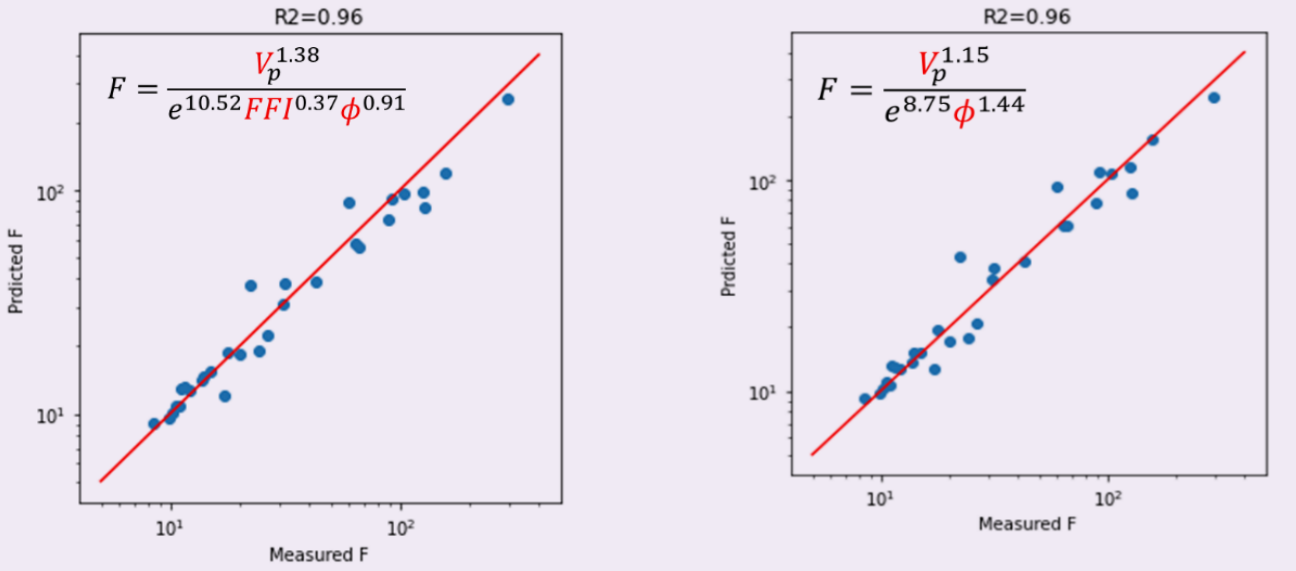
$$F = \frac{V_p^{1.15}}{e^{8.75} \phi^{1.44}}, \quad (14)$$

respectively, with their corresponding performance, Fig. 8.

In the SR training process, we have selected V_p, *FFI*, ϕ_{macro} and ϕ, as candidate input variables, based on the information in the correlation heat maps, Fig. 5, for all instances of training runs. Yet, the resultant equations may not include all of these variables. The SR algorithm determines which variables are more relevant. The embedded algorithm favors to avoid overcomplicated equation form, therefore, usually only a subsect of the entire selected candidate input variables may be included in the target equation. In this case, Eqn. 13 has three independent input variables: V_p, *FFI*, and ϕ, and Eqn. 14 only has two: V_p and ϕ. The difference is not surprising considering (a) the data set is small — the random selection of training data being included in training matters, and (b) the random nature of the genetic programming-based crossover and mutation processes also contributes to the diversity of the equations.

Despite the diversity in the prediction equation forms, these equations are equally valued in view that they pass the same training process, and when the data set are sufficiently large and diverse, the performance should have an insignificant difference. Even if the data set is small, the random training data allocation prohibits artificial

Fig. 8 *The performance of the two F prediction equations trained by the same database. The randomness of the genetic programming-based SR and the random selection of training data inclusion yield different prediction equations, equally valid, but may vary input variable dependences.*



bias, but a too small data set could result in unintentional data diversity fluctuation. In such a case, evaluation of the model performance comparison provides one additional layer of guard from performance fluctuation.

In the present case, both models perform equally well with $R^2 = 0.96$. Noting this R^2 value is identical that with Eqn. 9, we are more confident that either of these models are quite optimized.

k Prediction

The previous example is a relatively simple case as the data are well behaved in the sense that the correlation between input variables and target are strong. From case to case, we may encounter data that are more diverse, noisier, and/or weak correlations between inputs and target. To demonstrate that the SR can also handle such more challenging cases, we apply our workflow on the same core data set to obtain *k* prediction equations.

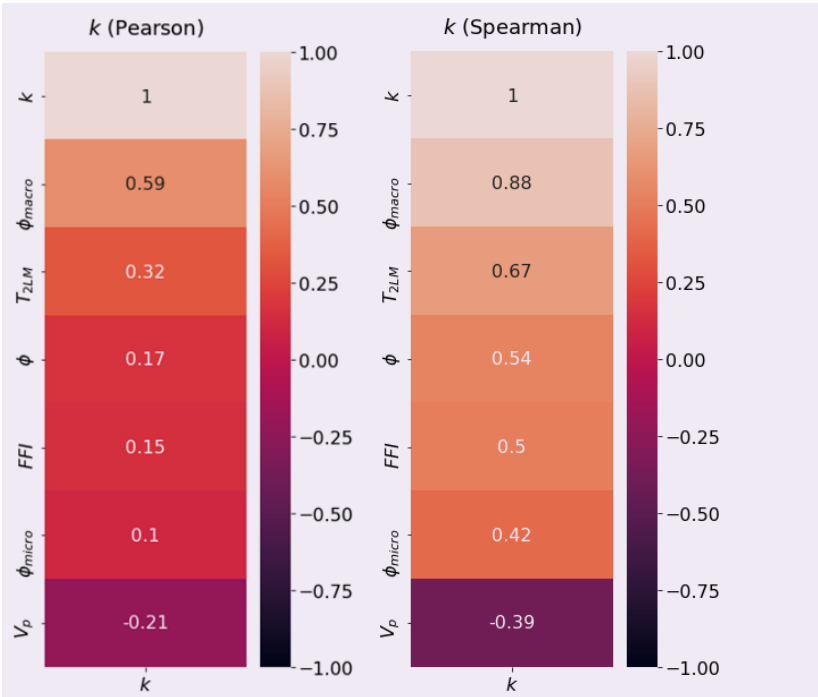
The heat maps for *k* were previously shown in Fig. 2, and the sorted heat maps are shown in Fig. 9.

It is quite obvious from Fig. 9, that the *k* does not linearly correlate to either of these candidate input variables strongly. It is somewhat better correlated to a few candidates in the monotonic sense (Spearman). Clearly, the correlations are much weaker than that for *F*. For such weaker correlation cases, coupled with a small available data set, the prediction problem becomes challenging.

The Spearman heat map in Fig. 9 indicates that those candidate input variables that have relatively high correlation coefficients ($\rho > 0.5$) are all deliverables from NMR measurements. This suggests to us first to try using NMR deliverables alone to train the *k* model.

We conducted several SR training runs using DataRobot to obtain several equations with varying, mediocre performance. Figure 10 shows four of the results. The R^2 values vary from 0.47 to 0.69. Although all of them perform better than the commonly used NMR-based SDR *k* equation with default parameters (Fig. 11 with a $R^2 = -0.98$), none of them are highly satisfactory. Therefore, the next step is necessary to improve the model.

Fig. 9 *Sorted heat maps for the k model development.*



Multiple runs generated equations that show different

Fig. 10 *Four instances of training run results with a small data set of weak correlations for permeability equation with NMR deliverables.*

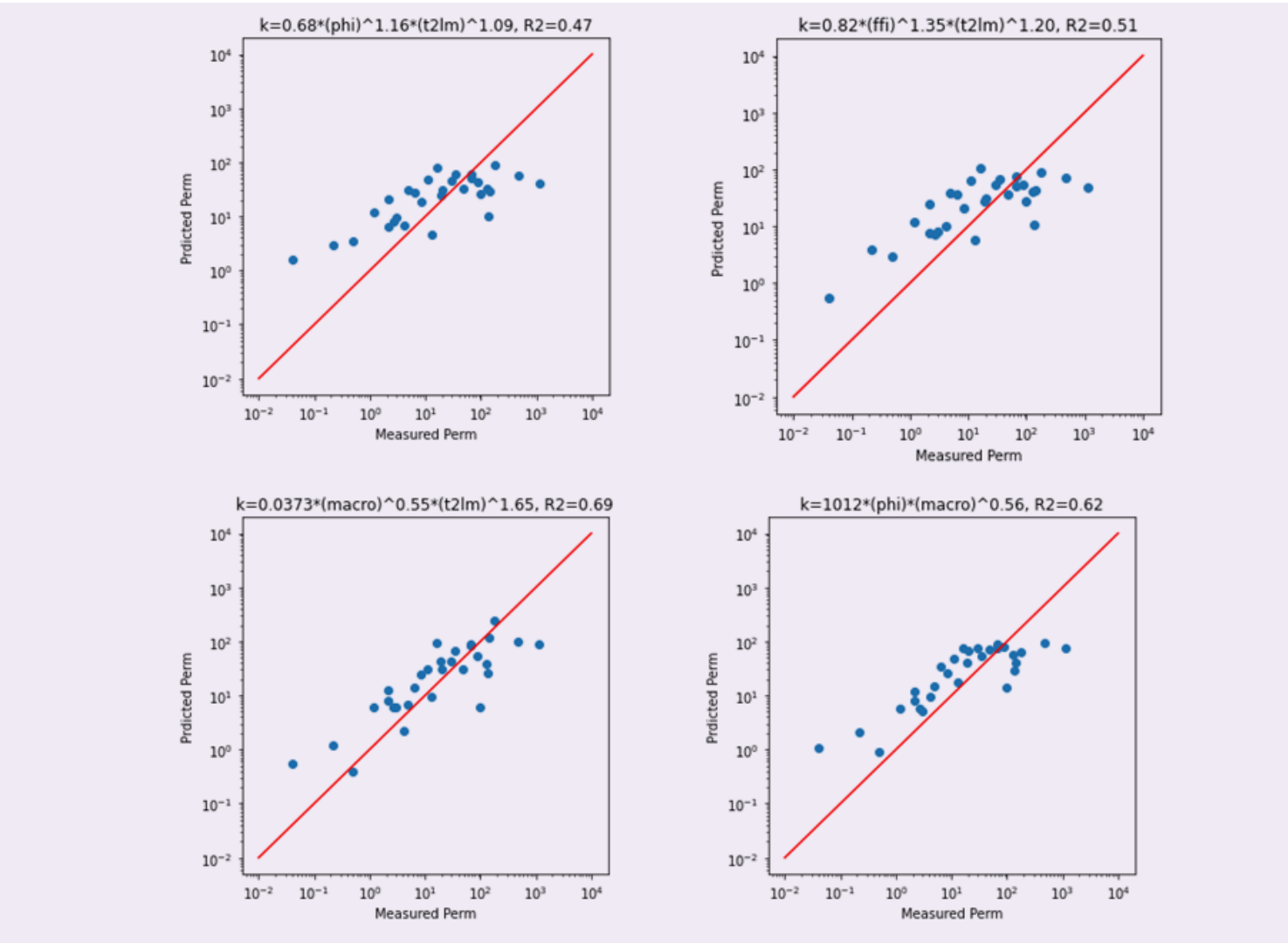
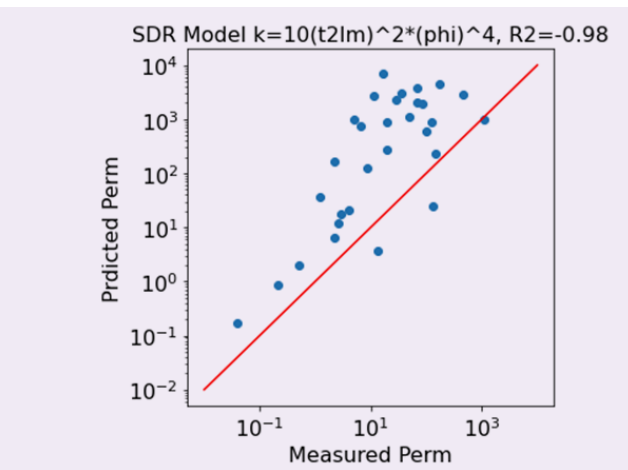


Fig. 11 *A comparison of the SDR* k *equation with default parameters calculated* k *vs. measured* k.



dependencies, five of which are shown in Eqns. 15 to 18, each containing a different set of input variables, even though the included candidate input variable sets for training are the same. The small number of input variables in each of the equations is understandable as the noisy, not so strongly correlated inputs and the target usually do favor fewer variables in the equation. The diversity of the dependencies again attributes to small training data sets randomly selected, and the random process in genetic programming based SR.

$$k = 0.68\phi^{1.16}T_{2LM}^{1.09} \quad (15)$$

$$k = 0.037\phi_{macro}^{0.55}T_{2LM}^{1.65} \quad (16)$$

$$k = 1012\phi \cdot \phi_{macro}^{0.56} \quad (17)$$

$$k = 0.037\phi_{macro}^{0.55}T_{2LM}^{1.65} \quad (18)$$

In this case, the ensemble modeling approach is applied by taking the logarithmic operation on each side of these equations first. Recognizing that all these dependencies and the function forms are valid, then the ensemble modeling equation should take the form of Eqn. 6:

$$\ln k = a + b1 \ln \phi + b2 \ln \phi_{macro} + b3 \ln T_{2LM} \quad 19$$

Then, we use this function form to refit all available training data (i.e., obtain a, $b1$, $b2$, and $b3$) to obtain:

$$k = 26.37\phi^{0.72}\phi_{macro}^{0.85}T_{2LM}^{0.76} \quad 20$$

Using this prediction equation, the performance is improved to $R^2 = 0.75$, better than any of the individual instances. Figure 12 shows a comparison of the predicted k and measured k.

While ensemble modeling visibly improves model performances, to this point, our attempts on this problem were initially limited in the function form selection. The simple model form is always worth the effort to try initially, as a baseline for comparing with more complicated models.

In the next attempt, we want to see if the model can improve visibly by inclusion of max and min operators, which allows conditional branching. A new equation is obtained as:

$$k = 1394\phi^{2.26}e^{0.95max\left[-2.19,\ 5.35+\ln\phi+3.98\ln\left(\frac{\phi_{macro}}{\phi}\right)\right]}. \quad 21$$

To understand whether the branching criteria max $[-2.19, 5.35 + \ln \phi + 3.98 \ln \frac{\phi_{micro}}{\phi}]$ makes any physical sense, we plotted the relationship between porosity and $\frac{\phi_{micro}}{\phi}$ in the manner shown in Fig. 13. The data satisfying $5.35 + \ln \phi + 3.98 \ln \frac{\phi_{micro}}{\phi} < -2.19$ are plotted in a yellow star symbol and those satisfying $5.35 + \ln \phi + 3.98 \ln \frac{\phi_{micro}}{\phi} > -2.19$ are plotted in a blue cross symbol. This plot shows that the yellow symbol represents samples whose macropore is only a small fraction of its total porosity, while the rest of the samples all have a larger fraction of macropores.

Fig. 12 *The ensemble model predicted* k *vs. measured* k.

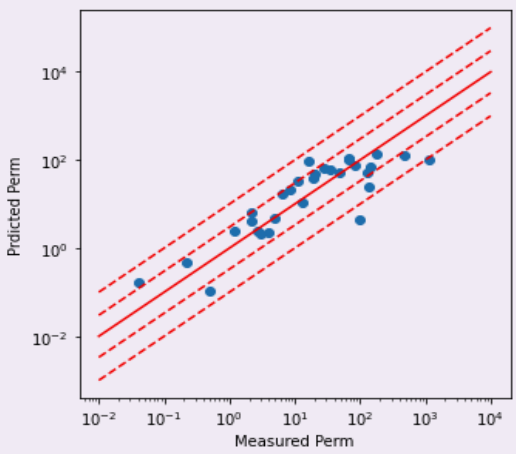
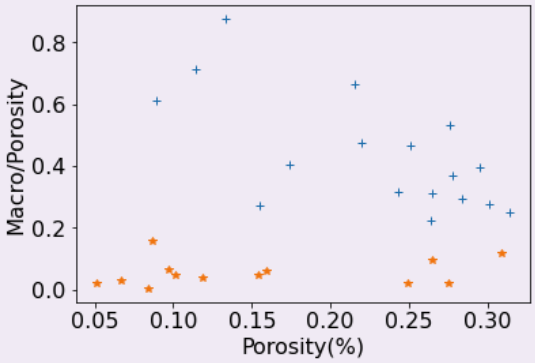


Fig. 13 *Examination of two key input variables for helping understand the physical meaning of the branching condition.*



Being able to examine the physical meaning of the SR delivered prediction equation certainly gives us a means to evaluate the rational of conditional branching. Such capability is absent from any black box AI methods. Recalling that Gomaa et al. (2006)[1] has developed an empirical k model from a similar rock type based on a large number of experimental data, they found two equations: Equation 22 for mesopore and micropore dominated carbonate rocks and Eqn. 23 for macropore dominated carbonates.

$$k = 0.5\phi^2(\rho T_{2LM})^2 \quad 22$$

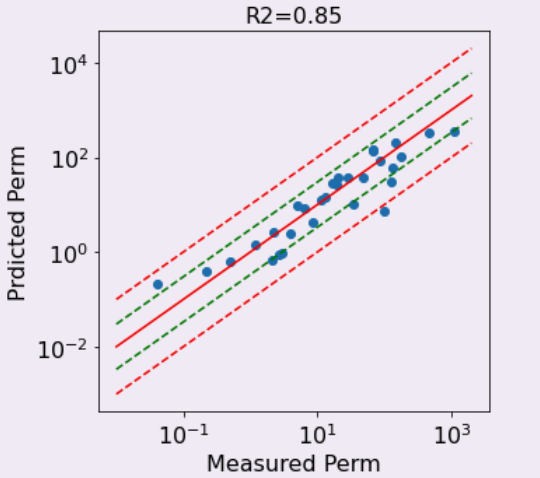
$$k = C\phi^2\left(\frac{\phi_{macro}}{\phi - \phi_{macro}}\right)^2 \quad 23$$

While these two equations differ in form from Eqn. 21 obtained with SR, the branching conditions have similar physical meaning. Empirically obtaining conditional branched equations usually is not a trivial task. It takes a lot of data and experience. In the case of Eqns. 22 and 23, the two forms of the equations resemble the two popularly used SDR and Coates[4] models, respectively, and branched condition is basically to find out which one of them is better applicable in which branching condition. On the other hand, using the SR method and obtaining a conditional branching equation does not require particular petrophysical experience, letting the data to speak for itself whether the branching is needed and what the expression is for the condition.

With Eqn. 21, the prediction performance is much improved to $R^2 = 0.85$, better than the ensemble modeling derived with Eqn. 20. Figure 14 shows the comparison of the predicted vs. the measured results.

Although Eqn. 21 already meets our expectation, in general, we would like to try various different ways of training to obtain multiple prediction equations for the same target parameter. These different ways of training may include a different subset of input variables, inclusive or exclusive of certain mathematical function forms,

Fig. 14 *The predicted* k *using conditional branching Eqn. 22 vs. measured* k.



and/or with different termination conditions. Two such attempts yield two new equations, respectively:

$$k = 1088\left(\frac{V_p}{5000}\right)^{7.43}\phi^{4.23}\exp\frac{-12.85}{\ln\phi_{macro}} \quad 24$$

and

$$k = 33.4\phi^{3.09}\exp\frac{8.8\ln\phi}{\ln\phi_{macro}}. \quad 25$$

The corresponding performance comparisons are shown in Fig. 15 and Fig. 16, respectively.

Equation 24 is obtained with the inclusion of the sonic measurement *Vp* in SR process. Equation 25 is obtained by including the exponential operator in the SR process. In the described applications of the SR-based model equation development process, several mathematically equally valid equations are created due to the random nature of genetic programming processes, the customized inclusion of initial elementary functions and operators. Naturally, one may raise the question of "Which one is optimal to use?" The answer is not always straightforward, therefore we apply the model discrimination process, previously described, to help us to decide.

The mathematical complexity for Eqns. 24 and 25 are Level 5 and Level 7, respectively. From a measurement complexity point of view, Eqn. 24 contains three input variables (three types of measurements: V_p, ϕ, and ϕ_{macro}), but Eqn. 25 contains two input variables (two types of measurements: ϕ and ϕ_{macro}). ϕ can be obtained from NMR, neutron, and density logging measurements, ϕ_{macro} can be obtained by NMR only, and V_p by sonic only.

Therefore, using Eqn. 24 requires at least inputs from two logging tools, but Eqn. 25 requires only NMR data. Consequently, the physics complexity level for Eqn. 25 is lower than that of Eqn. 24. This, together with a higher model performance ($R^2 = 0.84$ for Eqn. 25 vs. $R^2 = 0.75$ for Eqn. 24), the preference goes to Eqn. 25.

This model discrimination process potentially can be automatized by a defined physics complexity level and a quantitative score, based on a customized formula considering the cost of a particular type of logging measurements, data availability, and data reliability. Such quantitative scoring is likely to be varied for different operators. For instance, NMR logging is routinely acquired for carbonate reservoirs, but in another type of formation, it is not always available.

In this article, we only used R^2 for model performance comparison, while other error metrics can also be used

Fig. 15 *The predicted* k *by Eqn. 24 vs. measured* k.

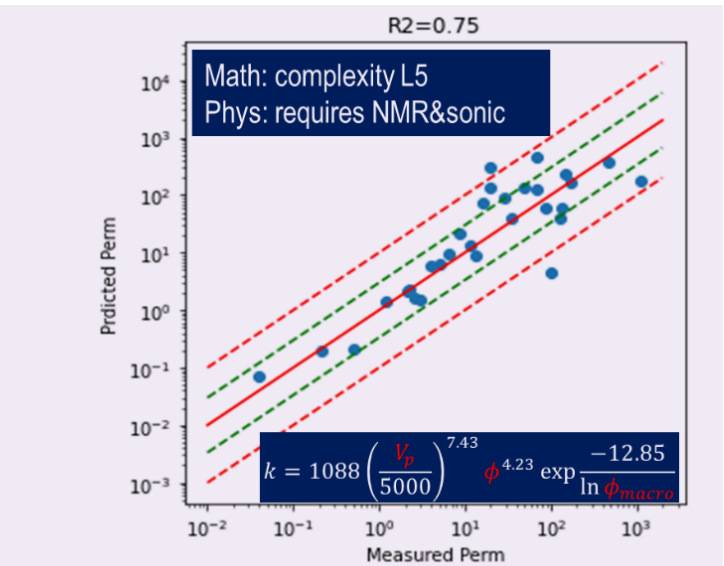
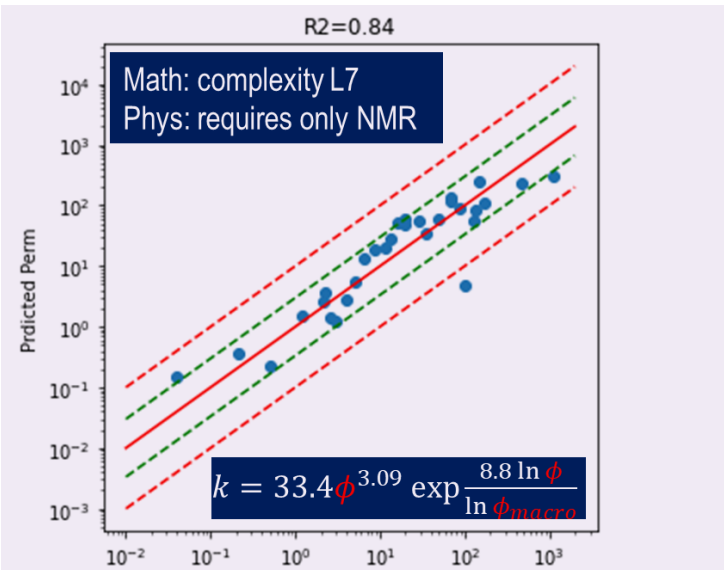


Fig. 16 *The predicted* k *by Eqn. 25 vs. measured* k.



to replace or used in addtion to R^2 as the performance measure(s).

Conclusions

In this article, we described a new, non-black box, AI-based petrophysical interpretation model development methodology, which consists of five components each for a distinguished purpose. Correlation heat maps are used to optimize the model input variables pool, and genetic programming-based SR algorithms are applied, which combines multiphysics measurement data objectively in creating the model though random processes. The delivered model is an equation allowing petrophysicists to evaluate the rational of such an equation. The ensemble modeling approach is incorporated in the workflow as an option, which is particularly useful when the training data set is small, and/or correlation between the input and the target is marginal.

By designing the primitive set of function forms and operators based on measurement physics knowledge and formation complexity, one may obtain a different level of complexity equations. To non-objectively determine the optimal model, a model discrimination framework is developed, which weighs collectively on equation complexity, measurement complexity, and model performance.

The new method has been applied to a carbonate data set for F and k prediction. The core measurement data used in the study are all logging deliverable equivalents. Therefore, even though for training purpose core data are used, the developed models are readily applicable to logging data interpretation.

The article went through in great detail why and how every component of the workflow are used. The application examples demonstrated that the SR-based method can be used to improve existing interpretation models by incorporating logging measurement to capture continuous variation of model parameters, as well as create new equations that are not constrained by the known equations, even with conditional branching.

Acknowledgments

This article was prepared for presentation at the SPWLA 63rd Annual Logging Symposium, Stavanger, Norway, June 10-15, 2022.

References

1. Gomaa, N.M., Al-Alyak, A., Ouzzane, D.E., Saif, O., et al.: "Case Study of Permeability, Vug Quantification, and Rock Typing for a Complete Carbonate," SPE paper 102888, presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, September 24-27, 2006.

2. Archie, G.E.: "The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics," *Transactions of the AIME*, Vol. 146, Issue 1, December 1942, pp. 54-62.

3. Kenyon, W.E., Day, P.I., Straley, C. and Willemsen, J.F.: "A Three-Part Study of NMR Longitudinal Relaxation Studies of Water Saturated Sandstones," *SPE Formation Evaluation*, Vol. 3, Issue 3, September 1988, pp. 662-636.

4. Coates, G.R., Miller, M., Gillen, M. and Henderson, C.: "The MRIL in Conoco 33-1 — An Investigation of a New Magnetic Resonance Imaging Log," paper presented at the 32nd SPWLA Annual Logging Symposium, Midland, Texas, June 16-19, 1991.

5. La Cava, W., Orzechowski, P., Burlacu, B., Olivetti, F., et al.: "Contemporary Symbolic Regression Methods and their Relative Performance," *arXiv:2107.14351v1 [cs.NE]*, July 2021.

6. Udrescu, S-M. and Tegmark, M.: "AI Feynman: A Physics-Inspired Method for Symbolic Regression," *Science Advances*, Vol. 6, Issue 16, April 2020.

7. Su, W., Cai, C., Wang, P.S., Li, H., et al.: "Complexity of Mathematical Expressions and Its Application in Automatic Answer Checking," *Symmetry*, Vol. 13, Issue 2, January 2021, pp. 188-203.

About the Authors

Dr. Songhua Chen

Ph.D. in Physics, University of Utah

Dr. Songhua Chen is Senior Manager of the Nuclear Magnetic Resonance (NMR) Sensor Physics at the Halliburton Houston Technology Center. Since joining Halliburton 11 years ago, he has been leading a team of scientists and mathematicians to research and design new wireline Xaminer Magnetic Resonance (XMR) and logging while drilling (LWD) NMR sensors, as well as to optimize NMR data acquisition, processing, and interpretation methodologies. Most recently, Songhua's interest focuses on carbonate pore typing, unconventional reservoir fluid identification, and applying data analytic approaches in petrophysical interpretation.

In the last several years, he has also been closely collaborating with colleagues in Aramco's RDD on a number of applied research projects to advance carbonate formation evaluation technologies, core analysis methods, and machine learning and deep learning applications.

Prior to joining Halliburton, Songhua was with Baker Hughes for 15 years as a Staff Scientist and NMR Interpretation Project Leader, and later became Senior Manager of the Integrated Interpretation group to develop technologies involving NMR, geochemistry, fluid sampling/testing interpretations, and pore scale modeling.

Prior to working in the energy service industry, he was a Senior Scientist at Texas A&M University in College Station, Texas.

Songhua is an inventor or co-inventor of more than 90 U.S. patents in NMR downhole sensors, data processing and analytics, inversion, core analysis, and integrated petrophysics. He has authored or coauthored more than 100 publications, including book chapters, peer-reviewed journal papers, and conference proceedings.

Songhua is an active member of the Society of Petroleum Engineers (SPE) and the Society of Petrophysicists and Well Log Analysts (SPWLA). He was previously a vice president of Publications of SPWLA and the editor of the *Petrophysics* journal. Songhua was selected twice as the SPWLA Distinguished Speaker in 2006 and 2013, respectively, and is the recipient of the 2019 SPWLA Distinguished Technical Achievement Award.

He received his B.S. degree from the Nanjing Institute of Technology, Nanjing, China, and his Ph.D. degree from the University of Utah, Salt Lake City, UT, both in Physics.

Dr. Wei Shao

Ph.D. in Applied Mathematics, University of South Carolina

Dr. Wei Shao is a Senior Scientific Advisor with Halliburton. Previously, he worked as a Senior Software Engineer with Baker Hughes. Wei has been primarily involved in developing algorithms and designing and implementing software applications for well logging data interpretation.

Wei received his B.S. degree in Applied Mathematics from Zhejiang University, Hangzhou, China, and his Ph.D. degree in Applied Mathematics from the University of South Carolina, Columbia, SC.

Dr. Huiwen Sheng

Ph.D. in Electrical Engineering, National University

Dr. Huiwen Sheng is a Senior Scientist of Signal Processing COE, at Halliburton Far East Pte. Ltd.

In 2016, Huiwen received her B.Eng. degree in Information and Communication Engineering (with honors) from Chu Kochen Honors College at Zhejiang University, Hangzhou, China, and in 2021, she received her Ph.D. degree in Electrical Engineering from the National University, Singapore, Singapore.

Dr. Hyung T. Kwak

Ph.D. in Physical Chemistry, Ohio State University

Dr. Hyung T. Kwak joined Saudi Aramco in April 2010 as a Petroleum Engineer with Saudi Aramco's Exploration and Petroleum Engineering Center – Advanced Research Center (EXPEC ARC). He had been a member of the Pore Scale Physics focus area and the SmartWater Flooding focus area of the Reservoir Engineering Technology Division. Since 2014, Hyung has been a focus area champion of the Pore Scale Physics focus area. His main research focus is seeking a deeper understanding of fluid-rock interaction in pore scale of the Kingdom's reservoirs.

Since joining Saudi Aramco, Hyung has been involved with various improved oil recovery and enhanced oil recovery (EOR) research projects, such as SmartWater Flooding, carbon dioxide EOR, and chemical EOR. Currently, he is leading a suite of key EXPEC ARC Fourth Industrial Revolution technology projects. Prior to joining Saudi Aramco, Hyung was a Research Scientist at Baker Hughes, from 2001 to 2010, focused on research related to nuclear magnetic resonance (NMR)/magnetic resonance imaging technology.

In 1996, Hyung received his B.S. degree in Chemistry from the University of Pittsburgh, Pittsburgh, PA, and in 2001, he received his Ph.D. degree in Physical Chemistry from Ohio State University, Columbus, OH.

Before moving into the oil and gas industry, Hyung was involved — as a postdoctoral fellow for 2 years — in a project developing the world's largest wide bore superconducting magnet NMR spectrometer, 900 MHz, at the National High Magnetic Field Laboratory.

He has more than 300 publications, including peer-reviewed articles and patents.

Fig. 8 *The matching quality of CNN, LSTM, BiLSTM, and CNN-LSTM in pressure prediction. Pressure training accuracy (a) and prediction accuracy (b) using CNN; training accuracy (c) and prediction accuracy (d) using LSTM; training accuracy (e) and prediction accuracy (f) using BiLSTM; and pressure training accuracy (g) and prediction accuracy (h) using CNN-LSTM.*

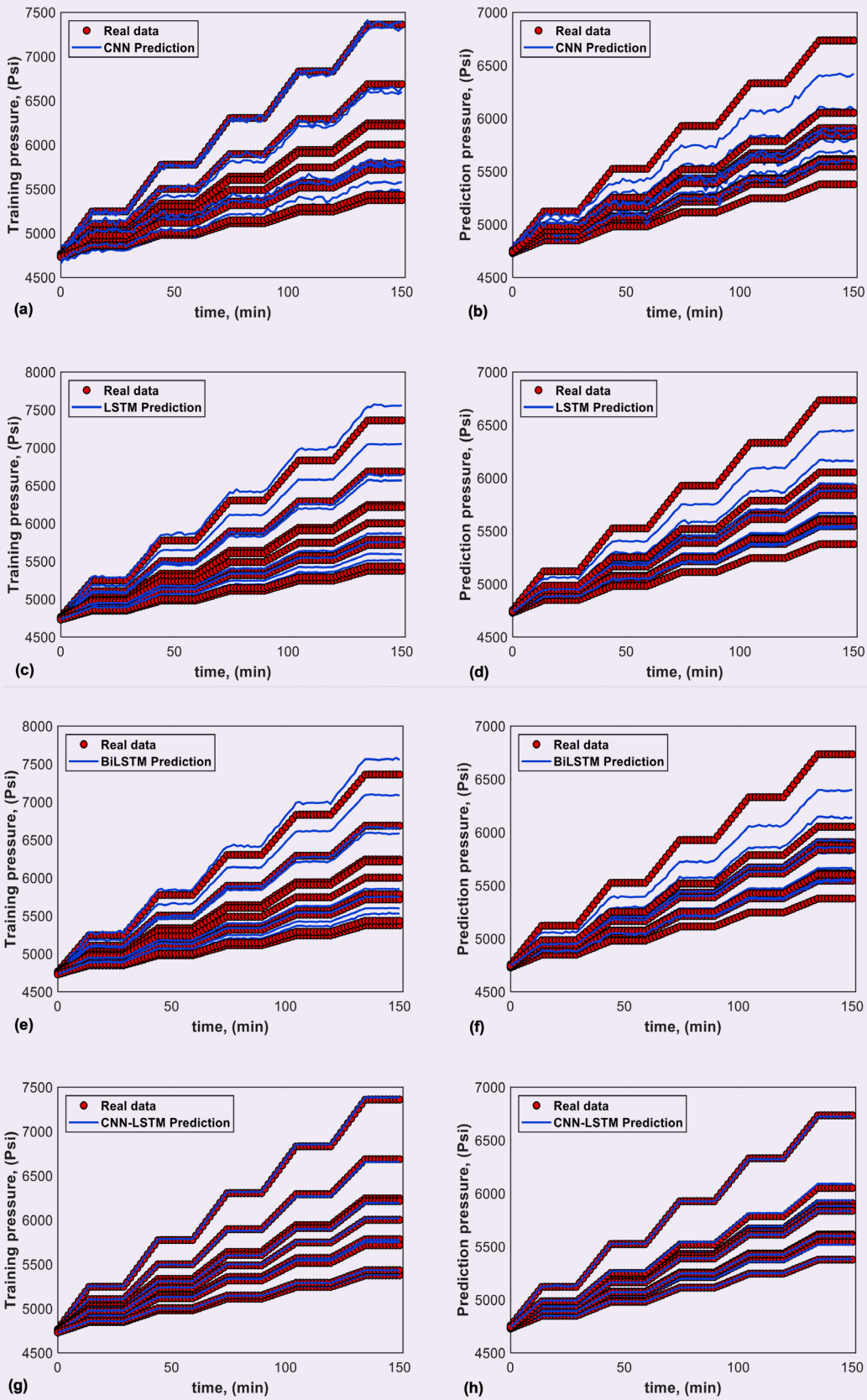
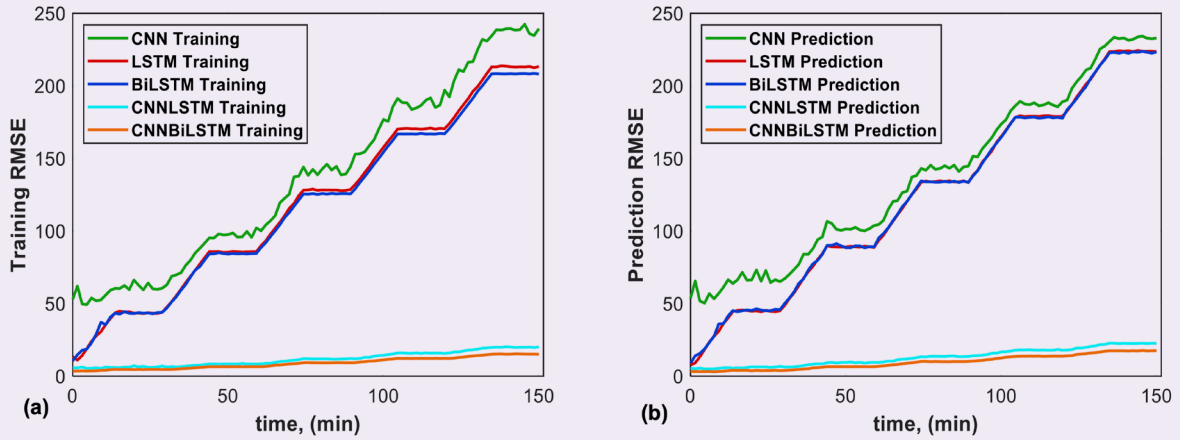


Fig. 9 *The performance of different surrogate methods toward (a) training RMSE, and (b) prediction RMSE.*



l m in the horizontal direction. Table l lists the variation of the parameters.

Figure 4 illustrates the inputs of the CNN-BiLSTM. One dynamic 2D image and two static 2D images are introduced to account for the time series and geological features, respectively. The pixel value in the injector location is the injection rate, and all the other pixel values are zero.

Figure 5a shows the pressure distribution after 90 minutes of CO_2 injection, and Fig. 5b shows the pressure change of the monitoring well, Well-F2, during the CO_2 injection. This case represents the typical behavior of the used model, which serves as the baseline case in this study.

Results

We use a synthetic CO_2 leakage case as the observation data, based on the base case as stated in Table l. We use the prediction from the CNN-BiLSTM model as the no leakage baseline. The objective is to demonstrate that the established model can detect CO_2 leakage.

First, we set the range of the uncertainty parameters. The CNN-BiLSTM is trained with 800 and tested with 100 cases generated using LHS. The Bayesian optimization is used to automate the tuning process of the hyperparameters, and we show the initial range and the optimized hyperparameters in Table 2.

The number of filters and the filter size in convolutional layers are set to 16 and 5, respectively. Figure 6a shows the matching accuracy of the training cases, and Fig. 6b shows the matching accuracy of the prediction accuracy. Figure 6c and Fig. 6d show the parity plots at t_D = 75 minutes and t_D = 150 minutes, respectively. The results show good matching for both the training and predictions.

We use a synthetic simulation model with the CO_2 leakage rate of 50 ft^3/d, and other parameters are set to the same as the base case to mimic the CO_2 leakage phenomenon. We can use the well-trained CNN-BiLSTM surrogate to predict the pressure change of Well-F2 in a no leakage situation, so the CNN-BiLSTM can serve as an anomaly detector of CO_2 leakage.

Figure 7a shows the anomaly detections, where the shaded zone represents the confidence intervals of no leakage occurred determine by the detection threshold σ. We can clearly see that the pressure anomalies tend to happen in the later stage of the pulse testing because more CO_2 has been leaked to the surface with increased testing time. Figure 7b compares the pressure change in the base case without CO_2 leakage, and the pressure change predicted using the CNN-BiLSTM. The matching shows good agreement and proves the effectiveness of the CNN-BiLSTM and the accuracy of the anomaly detections.

Comparison of Different Methods

We compare the performance of different surrogate models, such as CNN, LSTM, BiLSTM, and CNN-LSTM. Figures 8a to 8h shows the training and testing accuracy of the different proxy models. The training and testing set is set to 800 and 100, respectively. The range of the hyperparameters is the same as in Table 2, and the tuning process is automated using Bayesian optimization.

Figures 9a and 9b give the average training and testing root mean square error (RMSE). We can quickly notice that the CNN-BiLSTM achieves the highest accuracy, therefore, we use the CNN-BiLSTM as our proxy model.

Conclusions

We present a novel workflow to efficiently and robustly perform CO_2 leakage detection using CNN-BiLSTM and Bayesian optimization, where the CNN-BiLSTM is developed to capture both the time dependent and spatial information of the high fidelity model and reduce the computation time. Bayesian optimization is used to automate the tuning process. Specifically, the developed CNN-BiLSTM maps the nonlinear relationship between the input (permeability, porosity, and injection rate) and

the output (pressure change of Well-F2).

The use of the LHS is crucial because it produces space filling manner samples using a stratified sampling scheme, and therefore it provides the generality of the established proxy model. Other sampling methods usually generate clustering effects because they tend to generate unbalanced samples.

The pressure anomalies are mainly detected in the later stage of the injection period, which is caused by the accumulative leakage of the CO_2 during injection.

A detailed comparison of different proxies can be found in Fig. 9, where the CNN-BiLSTM has the lowest RMSE. The CNN cannot capture the time series feature and thereby performs the worst. The LSTM and BiLSTM cannot deal with spatial information, so the performance of the two surrogates is not satisfying. The CNN-LSTM can handle both the time series and spatial feature while performing relatively worse than the CNN-BiLSTM, since the CNN-BiLSTM can process the information bidirectionally.

The Bayesian optimization is used to automate the tuning task instead of the traditional trial-and-error process. Choosing an appropriate objective function when performing the Bayesian optimization is critical. In this work, the prediction RMSE is chosen as the objective function to maximize the predictability of the surrogate model.

Acknowledgments

We would like to thank Saudi Aramco for funding this research and CMG Ltd. for providing the IMEX academic license, KAUST for the support, and MATLAB for the software license.

This article was prepared for presentation at the SPE Annual Technical Conference and Exhibition, Houston, Texas, October 3-5, 2022.

References

1. Zhong, Z., Sun, A.Y., Yang, Q. and Ouyang, Q.: "A Deep Learning Approach to Anomaly Detection in Geological Carbon Sequestration Sites Using Pressure Measurements," *Journal of Hydrology*, Vol. 573, June 2019, pp. 885-894.

2. Sun, A.Y., Lu, J.M., Freifeld, B.M., Hovorka, S.D., et al.: "Using Pulse Testing for Leakage Detection in Carbon Storage Reservoirs: A Field Demonstration," *International Journal of Greenhouse Gas Control*, Vol. 46, March 2016, pp. 215-227.

3. Sinha, S., de Lima, R.P., Lin, Y., Sun, A.Y., et al.: "Normal or Abnormal? Machine Learning for the Leakage Detection in Carbon Sequestration Projects Using Pressure Field Data," *International Journal of Greenhouse Gas Control*, Vol. 103, December 2020.

4. Zhang, L., Huang, H., Wang, Y., Ren, B., et al.: "CO_2 Storage Safety and Leakage Monitoring in the CCS Demonstration Project of Jilin Oil Field, China," *Greenhouse Gases: Science and Technology*, Vol. 4, Issue 4, August 2014, pp. 425-439.

5. Das, A.K., Dejam, M. and Hassanzadeh, H.: "New Line-Source Solution and Scaling Relations for Diffusive Leakage of Brine from an Infinite Aquifer Caprock Composite Domain during Geological Storage of CO_2," *International Journal of Greenhouse Gas Control*, Vol. 118, July 2022.

6. March, R., Doster, F. and Geiger, S.: "Assessment of CO_2 Storage Potential in Naturally Fractured Reservoirs with Dual-Porosity Models," *Water Resources Research*, Vol. 54, Issue 3, March 2018, pp. 1650-1668.

7. Shao, Q., Boon, M., Youssef, A., Kurtev, K., et al.: "Modeling CO_2 Plume Spreading in Highly Heterogeneous Rocks with Anisotropic, Rate Dependent Saturation Functions: A Field Data Based Numeric Simulation Study of Otway," *International Journal of Greenhouse Gas Control*, Vol. 119, September 2022.

8. Qiao, T., Hoteit, H. and Fahs, M.: "Semi-Analytical Solution to Assess CO_2 Leakage in the Subsurface through Abandoned Wells," *Energies*, Vol. 14, Issue 9, April 2021.

9. Hochreiter, S. and Schmidhuber, J.: "Long Short-Term Memory," *Neural Computation*, Vol. 9, Issue 8, 1997, pp. 1735-1780.

10. Santoso, R., He, X., Alsinan, M.M., Kwak, H.T., et al.: "Bayesian Long-Short Term Memory for History Matching in Reservoir Simulations," SPE paper 203976, presented at the SPE Reservoir Simulation Conference, on-demand, October 26, 2021.

11. He, X., Zhu, W., Santoso, R., Alsinan, M.M., et al.: "CO_2 Leakage Rate Forecasting Using Optimized Deep Learning," SPE paper 206222, presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, September 21-23, 2021.

12. He, X., Santoso, R., Alsinan, M.M., Kwak, H.T., et al.: "Constructing Dual-Porosity Models from High-Resolution Discrete Fracture Models Using Deep Neural Networks," SPE paper 203901, presented at the SPE Reservoir Simulation Conference, on-demand, October 26, 2021.

13. He, X., Santoso, R. and Hoteit, H.: "Application of Machine Learning to Construct Equivalent Continuum Models from High-Resolution Discrete Fracture Models," IPTC paper 20040, presented at the International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia, January 13-15, 2020.

14. Rajabi, M.M., Javaran, M.H., Bah, A., Frey, G., et al.: "Analyzing the Efficiency and Robustness of Deep Convolutional Neural Networks for Modeling Natural Convection in Heterogeneous Porous Media," *International Journal of Heat Mass Transfer*, Vol. 183, Part B, February 2022.

15. Santoso, R., He, X., Alsinan, M.M., Kwak, H.T., et al.: "Uncertainty Quantification and Optimization of Deep Learning for Fracture Recognition," SPE paper 204863, prepared for presentation at the SPE Middle East Oil and Gas Show and Conference, November 28-December 1, 2021, event canceled.

16. Santoso, R., He, X. and Hoteit, H.: "Application of Machine Learning to Construct Simulation Models from High-Resolution Fractured Formation," SPE paper 197439, presented at the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, November 11-14, 2019.

17. He, X., Qiao, T., Santoso, R., Hoteit, H., et al.: "Gas Injection Optimization Under Uncertainty in Subsurface Reservoirs: An Integrated Machine Learning-Assisted Workflow," paper presented at the ARMA/DGS/SEG 2nd International Geomechanics Symposium, virtual, November 1-4, 2021.

18. He, X., Zhu, W., Santoso, R., Alsinan, M.M., et al.: "Fracture Permeability Estimation Under Complex Physics: A Data-Driven Model Using Machine Learning," SPE paper 206352, presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, September 21-23, 2021.

19. Albattat, R., He, X., AlSinan, M.M., Kwak, H.T., et al.: "Modeling Lost Circulation in Fractured Media Using Physics-Based Machine Learning," paper presented at the 83rd EAGE Annual Conference and Exhibition, Madrid, Spain, June 6-9, 2022.

20. He, X., Zhu, W., AlSinan, M.M., Kwak, H.T., et al.: "CO_2 Storage Capacity Prediction in Deep Saline Aquifers: Uncertainty and Global Sensitivity Analysis," IPTC paper 22463, presented at the International Petroleum Technology Conference, Riyadh, Kingdom of Saudi Arabia, February 21-23, 2022.

21. Alhubail, A., He, X., AlSinan, M.M., Kwak, H.T., et al.: "Extended Physics Informed Neural Networks for Solving Fluid Flow Problems in Highly Heterogeneous Media," IPTC paper 22163, presented at the International Petroleum Technology Conference, Riyadh, Kingdom of Saudi Arabia, February 21-23, 2022.

22. Yan, B., Harp, D.R., Chen, B., Hoteit, H., et al.: "A Gradient-Based Deep Neural Network Model for Simulating Multiphase Flow in Porous Media," *Journal of Computational Physics*, Vol. 463, August 2022.

23. Li, Y., He, X., Zhu, W., AlSinan, M.M., et al.: "Digital Rock Reconstruction Using Wasserstein GANs with Gradient Penalty," IPTC paper 21884, presented at the International Petroleum Technology Conference, Riyadh, Kingdom of Saudi Arabia, February 21-23, 2022.

24. Zhou, Z., Lin, Y., Wu, Y., Wang, Z., et al.: "Spatial-Temporal Densely Connected Convolutional Networks: An Application to CO_2 Leakage Detection," paper presented at the SEG International Exposition and Annual Meeting, Anaheim, California, October 14-19, 2018.

25. Chen, B., Harp, D.R., Lin, Y., Keating, E.H., et al.: "Geologic CO_2 Sequestration Monitoring Design: A Machine Learning and Uncertainty Quantification-Based Approach," *Applied Energy*, Vol. 225, September 2018, pp. 332-345.

26. Gundersen, K., Oleynik, A., Alendal, G., Skaug, H., et al.: "Ensuring Efficient and Robust Offshore Storage Use of Models and Machine Learning Techniques to Design Leak Detection Monitoring," paper presented at the 14th Greenhouse Gas Control Technologies Conference, Melbourne, Australia, October 21-26, 2018.

27. de Lima, R.P., Lin, Y. and Marfurt, K.J.: "Transforming Seismic Data into Pseudo-RGB Images to Predict CO_2 Leakage Using Prelearned Convolutional Neural Networks Weights," paper presented at the SEG International Exposition and Annual Meeting, San Antonio, Texas, September 15-20, 2019.

28. de Lima, R.P. and Lin, Y.: "Geophysical Data Integration and Machine Learning for Multi-Target Leakage Estimation in Geologic Carbon Sequestration," paper presented at the SEG International Exposition and Annual Meeting, San Antonio, Texas, September 15-20, 2019.

29. Chen, B., Mehana, M.Z. and Pawar, R.J.: "Toward Quantitative Approach to Evaluating Greenhouse Gas Leakage from CO_2 Enhanced Oil Recovery Fields," SPE paper 209440, presented at the SPE Improved Oil Recovery Conference, virtual, April 25-29, 2022.

30. Gan, C., Feng, Q. and Zhang, Z.: "Scalable Multi-Channel Dilated CNN-BiLSTM Model with Attention Mechanism for Chinese Textual Sentiment Analysis," *Future Generation Computer Systems*, Vol. 118, May 2021, pp. 297-309.

31. Rai, A., Shrivastava, A. and Jana, K.C.: "A CNN-BiLSTM Based Deep Learning Model for Mid-Term Solar Radiation Prediction," *International Transactions on Electrical Energy Systems*, Vol. 31, Issue 9, September 2021.

32. Challa, S.K., Kumar, A. and Semwal, V.B.: "A Multibranch CNN-BiLSTM Model for Human Activity Recognition Using Wearable Sensor Data," *The Visual Computer*, August 2021, pp. 1-15.

33. Soltanian, M.R., Amooie, M.A., Cole, D.R., Graham, D.E., et al.: "Simulating the Cranfield Geological Carbon Sequestration Project with High-Resolution Static Models and an Accurate Equation of State," *International Journal of Greenhouse Gas Control*, Vol. 54, Part 1, November 2016, pp. 282-296.

About the Authors

Dr. Xupeng He

Ph.D. in Energy Resources and Petroleum Engineering, King Abdullah University of Science and Technology

Dr. Xupeng He works as a Postgraduate Researcher at the Ali I. Al-Naimi Petroleum Engineering Research Center, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

Xupeng's research interests include modeling fluid flow in fractured rocks, data-driven surrogate model development using machine learning techniques, uncertainty quantification, sensitivity analysis, and multi-objective optimization for subsurface problems.

He was the recipient of the Dean's Award in 2022 at KAUST. Also, Xupeng received the National Scholarship from the Ministry of Education of the People's Republic of China in 2012 and 2013.

Xupeng received his M.S. degree in Earth Science and Engineering and his Ph.D. degree in Energy Resources and Petroleum Engineering from KAUST in 2018 and 2022, respectively.

Marwah M. AlSinan

M.S. in Petroleum Engineering, Imperial College

Marwah M. AlSinan joined Saudi Aramco in 2013 as a Petroleum Engineer, working with the Reservoir Engineering Technology Division in the Exploration and Petroleum Engineering Center – Advanced Research Center (EXPEC ARC).

Her research interests include multiphase flow in fractures, carbon dioxide sequestration, and applications of nuclear magnetic resonance in porous media.

In 2013, Marwah received her B.S. degree in Petroleum and Natural Gas Engineering from Pennsylvania State University, State College, PA. She received her M.S. degree in Petroleum Engineering from Imperial College, London, U.K. in 2017.

Dr. Yiteng Li

Ph.D. in Earth Science and Engineering, King Abdullah University of Science and Technology

Dr. Yiteng Li works as a postdoc fellow at the Ali I. Al-Naimi Petroleum Engineering Research Center at the King Abdullah University of Science and Technology.

His research interest includes compositional simulation using novel flash techniques, pore-scale nuclear magnetic resonance relaxation simulation, digital rock physics, and machine learning.

Yiteng received his M.S. degree in Petroleum Engineering from the University of Southern California, Los Angeles, CA, and his Ph.D. degree in Earth Science and Engineering from King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

Dr. Hyung T. Kwak

Ph.D. in Physical Chemistry, Ohio State University

Dr. Hyung T. Kwak joined Saudi Aramco in April 2010 as a Petroleum Engineer with Saudi Aramco's Exploration and Petroleum Engineering Center – Advanced Research Center (EXPEC ARC). He had been a member of the Pore Scale Physics focus area and the SmartWater Flooding focus area of the Reservoir Engineering Technology Division. Since 2014, Hyung has been a focus area champion of the Pore Scale Physics focus area. His main research focus is seeking a deeper understanding of fluid-rock interaction in pore scale of the Kingdom's reservoirs.

Since joining Saudi Aramco, Hyung has been involved with various improved oil recovery and enhanced oil recovery (EOR) research projects, such as SmartWater Flooding, carbon dioxide EOR, and chemical EOR. Currently, he is leading a suite of key EXPEC ARC Fourth Industrial Revolution technology projects. Prior to joining Saudi Aramco, Hyung was a Research Scientist at Baker Hughes, from 2001 to 2010, focused on research related to nuclear magnetic resonance (NMR)/magnetic resonance imaging technology.

In 1996, Hyung received his B.S. degree in Chemistry from the University of Pittsburgh, Pittsburgh, PA, and in 2001, he received his Ph.D. degree in Physical Chemistry from Ohio State University, Columbus, OH.

Before moving into the oil and gas industry, Hyung was involved — as a postdoctoral fellow for 2 years — in a project developing the world's largest wide bore superconducting magnet NMR spectrometer, 900 MHz, at the National High Magnetic Field Laboratory.

He has more than 300 publications, including peer-reviewed articles and patents.

Dr. Hussein Hoteit

Ph.D. in Applied Mathematics, University of Rennes 1

Dr. Hussein Hoteit is an Associate Professor in Reservoir Engineering and the Program Chair of Energy Resources and Petroleum Engineering (ERPE) at King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

Before joining KAUST, Hussein worked for ConocoPhillips and Chevron Companies for about 12 years, where he conducted projects related to chemical enhanced oil recovery (EOR), CO_2 EOR, steam flood, EM heating, to name a few.

Hussein's current research includes chemical EOR, geological CO_2 storage, improved oil recovery optimization, data-driven machine learning, and reservoir simulation development.

He has earned several Society of Petroleum Engineers (SPE) awards, including SPE Distinguished Lecturer in 2009, and served as Associate Editor for the *SPE Journal* for more than 10 years.

Hussein received his B.S. degree in Pure Mathematics and Computer Sciences from Lebanese University, Lebanon, M.S. and Ph.D. degrees in Applied Mathematics from the University of Rennes 1, Rennes, France.