Arabian Journal of Geosciences (2021) 14: 545
https://doi.org/10.1007/s12517-021-06559-9

ORIGINAL PAPER

Auto-characterization of naturally fractured reservoirs drilled by horizontal well using multi-output least squares support vector regression

Seyedeh Raha Moosavi[1] · Behzad Vaferi[2] · David A. Wood[3]

Received: 28 September 2019 / Accepted: 16 January 2021 / Published online: 18 March 2021

Abstract

Pressure transient response (PTR) of horizontal well in naturally fractured reservoirs (NFR) has a particular characteristic shape. This PTR is often used to estimate parameters of NFRs and detect their wellbore and boundary regimes. Interporosity flow coefficient (λ) and storativity ratio (ω) are two important parameters of the NFR that often estimated by matching process on the PTR. Since the matching techniques' results are not often unique, in this study, the multi-output least squares support vector regression (MLS-SVR) is employed for simultaneous estimation of λ and ω. A databank of 500 PTRs for horizontal wells in naturally fractured reservoirs is generated by the finite element method, converted to the pressure derivative (PD) curves, and then used to develop and evaluate this auto-characterization paradigm. The predictive accuracy of the model is checked and validated by both smooth and noisy PTRs. The proposed model predicts ω and λ with overall absolute average relative deviations (AARD) of 0.186% and 3.754%, respectively. The correlation coefficients (R^2) of 1 and 0.99992 are obtained for the prediction of ω and λ, respectively. The Leverage outlier detection technique justified that only less than 6% of the predictions are within the suspect region. This MLS-SVR model can be simply integrated with commercial pressure transient analysis (PTA) packages for accurate prediction of ω and λ even from the noisy PTRs.

Keywords Horizontal wells · Naturally fractured reservoirs · Storativity ratio · Interporosity flow coefficient · MLS-SVR

Introduction

Pressure transient signals of underground porous media, including oil, gas, gas condensate, and water reservoirs, are valuable information sources. Despite the high importance of pressure transient response (PTRs), the characterization of fluid flow in porous media from the PTR is challenging. Like all other actual signals, the PTR is often poisoned by noisy data that are difficult to either be distinguished or removed (Moosavi et al. 2018a; Su et al. 2019; Chen et al. 2021). The primary objectives of reservoir fluid flow characterization are to improve the reservoir modeling and simulation accuracy, enhance oil recovery, provide production-performance predictions of different reservoir scenarios, and help determine remaining recoverable oil reserves (Cheng et al. 2016). Generic reservoir characterization utilizes multiple data sources, including well logging, PTR, geology, core analysis, and seismic. Such information helps to elucidate the drive mechanisms, reservoir and wellbore geometry, fluid types, rock properties, saturations, and dominant flow directions (Landa et al. 2000; Abdel-Fattah et al. 2015; Khan et al. 2016; Ciftci 2018; Yang et al. 2020). Some researchers focused on different aspects of facilities used to transport oil and gas from the reservoir to consumers (He et al. 2018; Liu et al. 2019b). A multi-dimensional mixed-integer nonlinear optimization model based on modified particle swarm optimization (MPSO) algorithm is proposed for a large-scale oil and gas gathering system (Liu et al. 2019b).

Responsible Editor: Santanu Banerjee

✉ Behzad Vaferi
vaferi@iaushiraz.ac.ir; behzad.vaferi@gmail.com

[1] Department of Chemical and Petroleum Engineering, Shiraz University, Shiraz, Iran

[2] Department of Advanced Calculations, Chemical, Petroleum, and Polymer Engineering Research Center, Shiraz Branch, Islamic Azad University, Shiraz, Iran

[3] DWA Energy Limited, Lincoln, UK

Pressure transient signals make a key contribution to overall reservoir characterization.

This signal typically represents the reservoir pressure drop during changes the flow rates over specified periods. The PTR is long established as one of the best techniques for predicting well and reservoir properties by matching observed pressure responses with some ideal reservoir flow models (Huang et al. 2018; Nategh et al. 2019).

Naturally and hydraulic fracture reservoirs pose some additional challenges of the uneven distribution and orientation of their fractures and typically more significant heterogeneity than homogeneous reservoirs (Li et al. 2020). Consequently, although such reservoirs are known to contain significant in-place oil resources worldwide (Saidi 1983), their oil recovery factors tend to be small. This circumstance highlights the need for better reservoir characterization to improve fluid flow simulation models leading to more efficient drainage of these complex reservoirs. Variation of fracture propagation by the in situ stress, natural fissure development, temporary plugging, and treatment parameters was studied (Guo et al. 2020b). Fluid flow theory was first applied to naturally fractured reservoirs (NFR) by two research groups (Barenblatt and Zheltov 1960; Warren and Root 1963). Their studies develop a dual-porosity model for fractured reservoirs with two different key parts, i.e., a fracture with high fluid transmissibility and low storage capacity, and the matrix with low fluid transmissibility and high storage capacity. These specific characteristics of dual-porosity reservoirs can be quantified by deriving two useful metrics: interporosity flow coefficient (λ) and storativity ratio (ω). It is an earlier measure interaction between matrix and fractured sections, while the latter represents an amount of fluid stored in the fractured section relative to the total fluid held in the whole reservoir (Egya et al. 2017).

Although the different values of λ and ω produce no unique characteristic shapes in the PTR, they produce entirely different patterns in the pressure derivative (PD) graphs. This feature of pressure derivative plots makes it possible to apply various machine learning algorithms to estimate λ and ω using the characteristic shapes of such plots. Similar to wide applications of machine learning techniques in different fields of science and engineering (Zhang et al. 2018; Yang et al. 2019; Shi et al. 2020), they are also employed to predict reservoir underlying models (Vaferi et al. 2011, 2016; Ghaffarian et al. 2014), while some researches focused on predicting specific values for key reservoir parameters (Eslamloueyan et al. 2010; Şahin and Çiftçi 2016).

Most models proposed to predict reservoir fluid flow metrics using machine learning algorithms have applied artificial neural networks (ANN). Alajmi and Ertekin developed an ANN model applied to NFR (Alajmi and Ertekin 2007). Their model involved a 4th-degree polynomial fit to the semi-log relationship between well-test pressure and time data derived via a simulation model (Sierra 1986). Well-test data combined with rock properties and fluid compositions are considered the input variables for their ANN model. Well-test pressure derivative graphs show much greater sensitivity to λ and ω variability for dual-porosity NFR. Eslamloueyan et al. developed their ANN model so that the digitized pressure derivative graph roles as input for characterization of NFR system through prediction both λ and ω (Eslamloueyan et al. 2010). Deep convolutional neural network (Xu et al. 2018; Li et al. 2019; Chen et al. 2020a; Lv and Qiao 2020; Qian et al. 2020a, b) as a new generation of artificial intelligence techniques is recently applied for reservoir characterization from well-testing data (Daolun et al. 2020; Liu et al. 2020).

This study's main objective is to develop a powerful smart strategy, named multi-output least squares support vector regression (MLS-SVR), for simultaneous prediction of λ and ω from digitized PD graphs. For this purpose, a dataset including 500 simulated pressure-time patterns is generated for a NFR system drilled by a horizontal well with the infinite acting condition. This dataset is converted to PD graphs and then used to train a multi-output least squares support vector regression model to predict λ and ω. The proposed method does not rely upon any graphical or correlation between variables and requires no prior information about the NFR system and its associated flow regimes. Both simplicity and accuracy of the designed MLS-SVR model are among key advantages of this technique. It saves time in analyzing of pressure response of NFR systems drilled by the horizontal well.

Methods

Pressure transient responses of NFR

In a drawdown test, the flow rate is held approximately constant while the bottom-hole pressure is continuously measured during the production period. Figure 1 depicts a typical pressure transient curve for a NFR system drilled with a horizontal well for a drawdown test. This pressure transient data is generated through a solution of the governing equation for the NFR drilled with a horizontal well by the finite element method. The NFR governing equations have been developed using the Warren-Root approach with some basic assumptions applied, including the uniform thickness of the reservoir with impermeable lower and upper boundaries, radial flow, isothermal, single-phase, and slightly compressible fluid with constant viscosity and rock properties. In this model, only fracture-fed wells are considered, and it is further assumed that each continuum (fracture and matrix) is homogenous and separate. Moreover, as mentioned, constant rock and fluid properties are also assumed.

Radial change in the reservoir pressure results in a fluid movement towards the production well. During the early

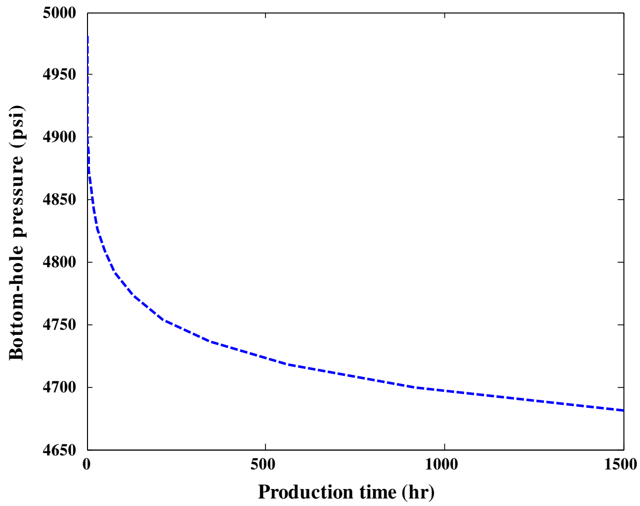

Fig. 1 Schematic of variation of pressure a horizontal well drilled in a naturally fractured reservoir during production period

production period, the bottom-hole pressure changes rapidly due to wellbore storage effects and the limited radius of investigation. Subsequently, as time progresses and the larger volume of the reservoir contributes to production, the bottom-hole pressure changes more slowly. Finally, the bottom-hole pressure represents the behavior of the outer boundaries.

Naturally fractured reservoirs

Naturally fractured reservoirs have two key parameters more than homogeneous ones. Therefore, it is useful to detail how these two parameters, i.e., interporosity flow coefficient and storativity ratio, influence on characteristic shapes of the NFRs. It is evident that they have an essential role in distinguishing pressure transient responses of the NRF from homogeneous reservoirs.

Storativity ratio

The storativity ratio is defined as the fraction of the total pore volume associated with one of the two porosities in a dual-porosity reservoir. Specifically, in NRF, ω refers to the volume fraction of total reserves contained within the fractures. It is defined as follows (Moosavi et al. 2018b):

$$\omega = \frac{(\varphi V c_t)_f}{(\varphi V c_t)_f + (\varphi V c_t)_m} \tag{1}$$

where φ and c represent porosity and compressibility factor, respectively. Subscript f and m stand for fracture and matrix, respectively.

The storativity ratio has a significant effect on the short-term fluid production deliverability of a reservoir. The ratio is typically between 0.01 and 0.1, and dual-porosity reservoir analysis is used in conjunction with the interporosity flow coefficient.

Interporosity flow coefficient

Interporosity flow coefficient is directly correlated with the permeability of the matrix segment of the reservoir, but inversely correlated with the permeability of the fractured segment of the reservoir. For dual-porosity reservoirs, λ typically exists in the value range 10^{-4} to 10^{-8}. This parameter can be mathematically expressed by Eq. (2).

$$\lambda = \alpha \frac{k_m}{k_f} L^2 \tag{2}$$

where k is the permeability, α shows the interporosity shape factor related to the block's matrix of the NFR, and L stands for half-length of the horizontal well.

Semi-log graph for NFR

As Fig. 2 shows, the dual-porosity components of a NFR generate two parallel straight lines on a semi-log graph. The semi-log is made up in the detail of three distinct components: (1) transient radial flow period, an early straight line component representing the homogeneous flow of the fractured segment of the reservoir before the matrix segment of the reservoir makes a noticeable contribution, (2) interporosity flow onset period—a transition component joining the two straight line components reflecting the first noticeable contribution from the matrix segment of the reservoir, and (3) composite homogeneous flow period—a later stage straight line component reflecting the combined contributions from the fractured segment and the matrix segment, beginning when fluids in the fracture segment are consistently replenished by fluids flowing from the matrix segment of the reservoir.

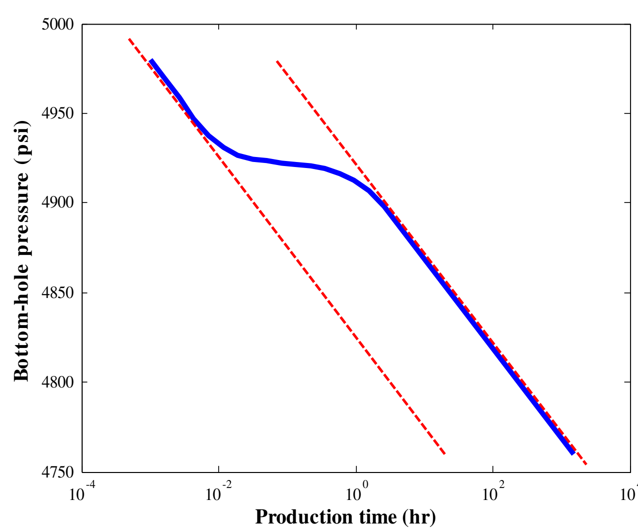

Fig. 2 Drawdown semi-log plot for a horizontal well drilled in a NFR

Effects of variation of interporosity flow coefficient and storativity ratio on a semi-log plot of a NFR drilled by a horizontal well are shown in Figs. 3 and 4, respectively. In Fig. 3, it is noticeable that as the interporosity flow coefficient increases, the first straight line component becomes shorter, and the transition component occurs sooner. Since the second zone on a semi-log plot (i.e., the transition component) involves some fluid flow from the matrix segment to the fractured segment of the NFR, any decrease in the permeability of the fractured segment or any increase in the permeability of the matrix segment will cause the fracture storage contribution to flow to be depleted more rapidly and the interporosity flow onset period to commence earlier.

As the storativity ratio increases, the first straight line component becomes shorter, the transition component commences earlier, and the transition to the composite homogeneous flow period also commences earlier.

Traditional methods for estimation of ω and λ

Mathematical models are widely used to interpret PTR to establish reservoir characterization purposes. A traditional approach employs semi-log and log-log plots to provide a detailed analysis of pressure transient information. To establish meaningfully accurate reservoir interpretations, these traditional methods need to be constrained in certain ways. For example, distinct flow regimes applicable to specific time intervals must be clearly delineated on both pressure and pressure derivative plots. If the flow regimes applicable to the specific time elapsed intervals are unknown, type curve matching methods fail to provide unique solutions. In such situations, a trial-and-error approach is required to apply various flow regimes to find the best fits to the semi-log curves. The direct synthesis technique avoids type curve matching to derive reservoir metrics from pressure transient data (Tiab 1989). It simultaneously uses pressure and PD curves to analyze vertically (Tiab 1994) and naturally fractured reservoirs (Engler and Tiab 1996) by exploiting analytical and empirical correlations.

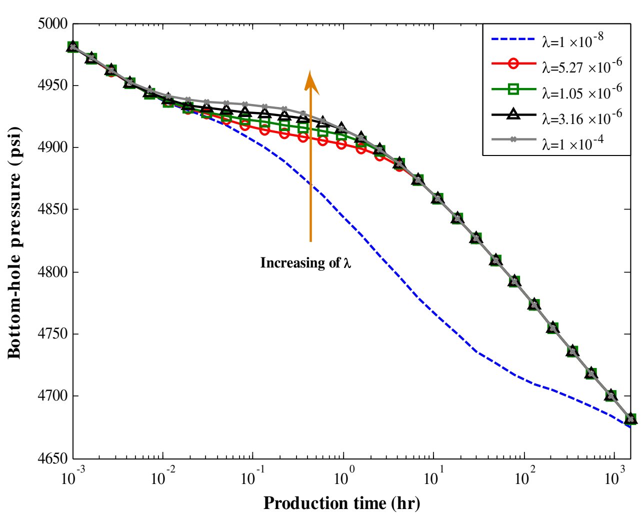

Fig. 3 Effect of interporosity flow coefficient on semi-log response of a horizontal well in a NFR (ω=0.055)

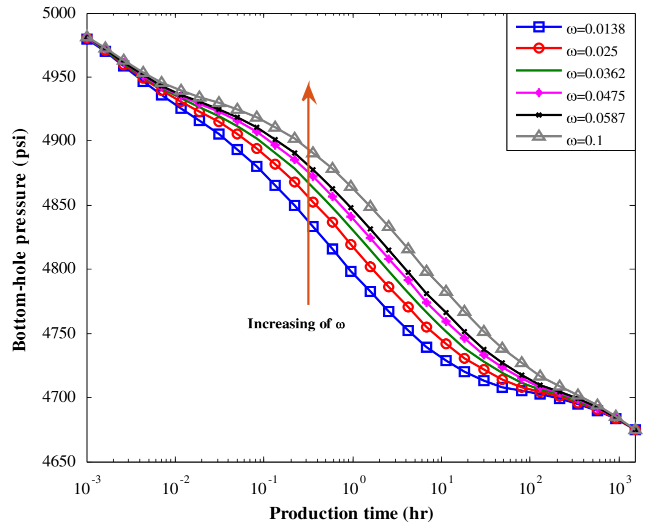

Fig. 4 Effect of storativity ratio on semi-log response of a horizontal well in a NFR (λ=1 × 10^{-8})

Pressure derivative for NFR system

Remarkably, the pressure derivative technique is a powerful and well-established interpretation tool in providing meaningful analysis of pressure transient signals (Bourdet et al. 1989; Escobar et al. 2018). As expressed by Eq. (3), this technique uses three points (i.e., pressure drop versus superposition time) for calculation of pressure derivative at a given point.

$$(PD)_k = \frac{\frac{(\Delta p)_k-(\Delta p)_{k-1}}{\ln(\Delta t)_k-\ln(\Delta t)_{k-1}}\left[\ln(\Delta t)_{k+1}-\ln(\Delta t)_k\right]+\frac{(\Delta p)_{k+1}-(\Delta p)_k}{\ln(\Delta t)_{k+1}-\ln(\Delta t)_k}\left[\ln(\Delta t)_k-\ln(\Delta t)_{k-1}\right]}{\ln(\Delta t)_{k+1}-\ln(\Delta t)_{k-1}} \tag{3}$$

These PD plots are more useful for parameter estimation and model detection than traditional ones (Eslamloueyan et al. 2010; Vaferi et al. 2011; Wang 2016). The PD technique specifies the real interpretation model of fluid using pattern matching of observed signal with some standardized type curves. After that, it is possible to quantify reservoir properties' values (Bourdet et al. 1989; Bourdet 2002; Tiab and Donaldson 2015).

Theoretically, the dual-porosity behavior of NFR usually appears as three distinct flow regimes in pressure derivative graphs: (1) transient radial flow period, an early straight line component associated with production from the fractured segment of the reservoir, (2) interporosity flow onset period, commencing in the middle period of the test representing the first noticeable production contribution from the matrix segment of the reservoir into the fractured segment, and (3) composite homogeneous flow period, the final test period during which simultaneous production from matrix segment and the fractured segment of the reservoir both contribute. The transition phase from the first flow regime to the second one appears as a distinct hump on the pressure derivative curve.

Figure 5 highlights the effect of changing the storativity ratio on the pressure derivative curves of NFR having horizontal well. As Fig. 5 clearly shows, by decreasing the storativity ratio, the first flow region finishes earlier. The transitional hump between first and second flow regimes moves to the left on the log-log plot. Also, the trough representing part of flow regime 2 progressively appears earlier as ω decreases. This results in the third flow regime commencing earlier as ω decreases.

Figure 6 presents the effect of variations in λ on the characteristic shape of the pressure derivative curve of NFR system. The earliest hump to appear in the pressure derivative curve is related to production from the fractured segment of the NFR. A decrease in λ can be induced by either decreasing the permeability of the matrix segment of the reservoir or increasing the permeability of the fractured segment of the reservoir. Such a decrease in λ will cause decreases in the pressure decline and the pressure derivative, which will lead to the earliest hump in the pressure derivative curve being more pronounced by extending over a more extended period. The second hump in the pressure derivative curve in Fig. 6 reflects fluid flow from the matrix segment to the fractured (i.e., third flow regime). An increase in λ can be induced by either decreasing the fracture permeability or increasing the matrix permeability. Such an increase in λ results in storage in the fractured segment of the reservoir being depleted earlier, leading to the third flow regime commencing earlier.

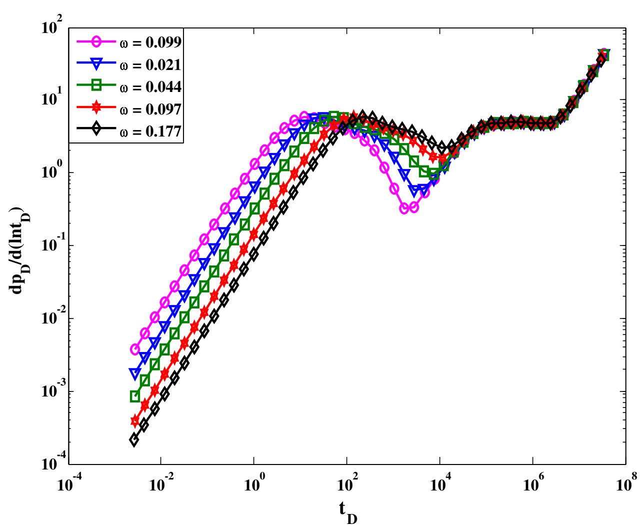

Fig. 5 Effects of variations of ω on the pressure derivative curve of a horizontal well in a naturally fractured reservoir

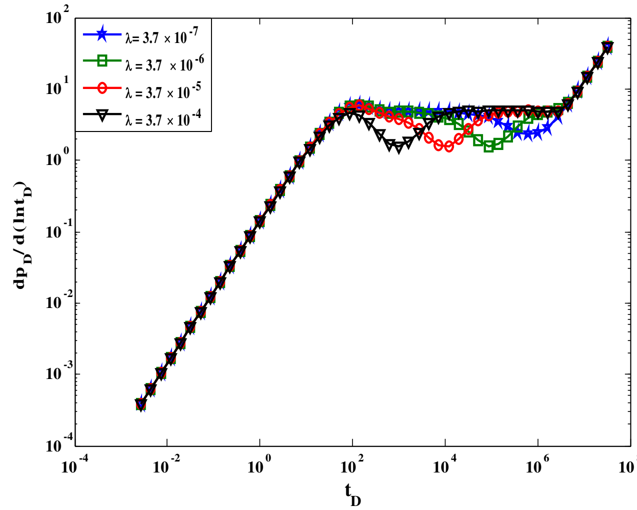

Fig. 6 Effects of variation of λ on pressure derivative plot of a horizontal well pressure-drawdown test in a NFR

Results and discussion

Interpretation model for NFR systems

In this study, the datasets, including 500 pressure transient signals for NFR system drilled by horizontal wells, are generated using an analytical solution of the governing equation of fluid flow using the finite element method. For generation these PTR, all reservoir properties except storativity ratio and interporosity flow coefficient are considered constant.

Table 1 Fluid and rock properties and wellbore geometry assumptions applied to the reservoir flow models evaluated in this study

Parameter	Value/ range	Unit
Well radius	0.3	ft
Formation thickness	40	ft
Porosity	0.17	–
Permeability	40	md
Vertical permeability	4	md
Skin factor	1	–
Viscosity	1.2	cp
Well length	800	ft
Total compressibility	5×10^{-6}	psi^{-1}
Oil formation volume factor	1.1	Rb/STB
Initial pressure	5000	psi
Simulation time	1500	hr
Top boundary	Sealing	–
Bottom boundary	Sealing	–
Interporosity flow coefficient	$10^{-4} - 10^{-8}$	–
Storativity ratio	$0.01 - 0.1$	–

The typical properties of NFR system for fluid, rock, and wellbore geometry are listed in Table 1.

Only the two-final metrics in Table 1, λ and ω, were allowed to vary while generating the PTR dataset. Although the general shape of all of the generated PTR (i.e., each with unique λ and ω values) are similar, their pressure derivative curves are quite distinct. Indeed, the MLS-SVR approach discriminates among these reservoir flow systems based on the difference in the characteristic shapes of their pressure derivative curves caused by interporosity flow coefficient and storativity ratio.

Accuracy of the MLS-SVR model

Three commonly used statistical measures of accuracy, i.e., correlation coefficients (R^2), absolute average relative deviations (AARD%), and mean square error (MSE), were used to the numerical evaluation of the performance of the developed MLS-SVR model. These statistical indices are mathematically expressed by Eqs. (4) through (6), respectively.

$$R^2 = \frac{\sum_{i=1}^{N}\left(PV_i^{act.} - \overline{\Delta PV}\right)^2 - \sum_{i=1}^{N}\left(PV_i^{act.} - \Delta PV_i^{cal.}\right)^2}{\sum_{i=1}^{N}\left(PV_i^{act.} - \overline{\Delta PV^{act.}}\right)^2} \tag{4}$$

$$AARD\% = \frac{100}{N}\sum_{i=1}^{N}\left(\left|\frac{PV_i^{act.} - PV_i^{cal.}}{PV_i^{act.}}\right|\right) \tag{5}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(PV_i^{act.} - PV_i^{cal.}\right)^2 \tag{6}$$

where N is the number of data records, $PV^{act.}$ and $PV^{cal.}$ indicate actual values of ω or λ and their associated predicted values by MLS-SVR model, and $\overline{\Delta PV}$ shows the average value of real data for ω and λ.

Table 2 displays the results of sensitivity analysis for the accuracy measures relating to ω and λ predictions for four different distributions of the training and testing subsets. All the divisions between training and testing subsets display high degrees of accuracy, but using the 90% of databank as training provided the best statistical indices for ω and λ in terms of AARD%, MSE, and R^2. The relatively large percentage errors for the AARD% relating to λ are due to the very low values (10^{-4} to 10^{-8}) of that metric. Indeed, even a very small difference between actual and calculated values for λ tends to produce a high level of error.

For a better presentation of these results and providing visual observation, the variation of the accuracy of the MLS-SVR models with training percent for both storativity ratio and interporosity flow coefficient is illustrated in Figs. 7 and 8, respectively.

Based on the results of sensitivity analysis presented in Table 2, Figs. 7 and 8, 90% (450 PD graphs) of the available databank is selected for the training of the MLS-SVR model.

Figures 9 and 10 depict actual values of ω and λ as a function of their associated predicted values by MLS-SVR model for the training and testing data subsets, respectively. The negligible difference between actual and predicted values confirms the highest degree of accuracy achieved by the MLS-SVR model for both storativity ratio and interporosity flow coefficient.

3.2. Performance of the developed MLS-SVR model for noisy signal

Unlike the generated PTR from a solution of governing equation which is typically smooth, signals associate with real systems typically has a level of noise and defects (Moosavi et al. 2020; Yue et al. 2020). Therefore, it is crucial to assess the performance of the MLS-SVR model when confronted with noisy data records. For this purpose, normal artificial noise was added to the ten different PD graphs of the NFR system drilled by a horizontal well. Figure 11 illustrates typical smooth as well as noisy pressure derivative records for a NFR flow regime with unique ω and λ values.

This figure clearly shows that a relatively high level of noise is added to the original curve. The developed MLS-SVR is provided with ten noisy PD graphs, and the obtained results are reported in Table 3. This table provides a comparison between actual values of ω and λ and their associated predicted values by the MLS-SVR model for noisy PD signals. This comparison reveals that noise in the data records does not significantly degrade the prediction performance of the MLS-SVR model, and it tolerates a relatively high level of uncertainty in independent variables.

Outlier detection

The presence of outlier and defect in the data records (Chen et al. 2020b; Guo et al. 2020a) making up the dataset influences the achievable accuracy by a model. As a large dataset has been generated and analyzed to develop the MLS-SVR model, some data record error is expected, resulting in some outliers in the predicted versus actual values. A useful statistical algorithm to detect outliers in datasets is the Leverage method enabling outlier points to be detected (Rousseeuw and Leroy 2005). This method determines standardized residuals between predicted and real values for each record in the dataset.

The distribution of standard residual has a mean of 0 and a standard deviation of 1. The leverage method then creates a Hat matrix for those standardized residuals defining a leverage index determined by Eq. (7):

Table 2 Sensitivity analysis results for proposed MLS-SVR model to training and testing subset divisions of the data records

Numbers of signals assigned to training/testing subsets	Dataset	Sensitivity analyses					
		Storativity ratio			Interporosity flow coefficient		
		AARD%	MSE	R^2	AARD%	MSE	R^2
375/125	Training	0.225	1.38×10^{-8}	0.99996	4.246	1.32×10^{-14}	0.99996
	Testing	0.399	2.71×10^{-8}	1	41.188	2.28×10^{-14}	0.99978
400/100	Training	0.243	1.50×10^{-8}	0.99996	5.318	1.38×10^{-14}	0.99996
	Testing	0.226	1.82×10^{-8}	0.99958	34.951	1.96×10^{-14}	0.99983
425/75	Training	0.182	9.11×10^{-9}	0.99998	2.191	6.71×10^{-15}	0.99998
	Testing	0.209	1.28×10^{-8}	1	32.034	8.06×10^{-15}	0.99984
450/50*	**Training**	**0.177**	8.81×10^{-9}	**0.99998**	**2.755**	6.37×10^{-15}	**0.99998**
	Testing	**0.265**	1.47×10^{-8}	**1**	**12.743**	1.10×10^{-14}	**0.99945**

* The most accurate MLS-SVR model. Note: multiple cases were executed with different signals assigned to the training and testing subsets

$$H = X\left(X^T X\right)^{-1} X^T \quad (7)$$

where X is a matrix of standardized residuals with $p \times q$ elements and X^t is its transpose. p and q represent the number of data records (500 in the NFR dataset generated) and the number of model parameters (30 data points used to define each pressure derivative curve), respectively.

The diagonal elements of the H matrix establish the leverage index. A William's diagram cross plotting the standardized residual and the leverage index calculated by Eq. (7) for each data record then reveals suspected outlying data records (Gramatica 2007).

Figures 12 and 13 display William's plots for ω and λ values obtained by the MLS-SVR model, respectively.

In these plots, the warning leverage value (H^*) is displayed as a vertical dashed line calculated using Eq. (8):

$$H^* = 3(n+1)/m \quad (8)$$

Here, n is the number of model parameters, m shows the number of data records. The cutoff value of 3 is typically applied in Eq. (8), representing a range of plus or minus three standard deviations from the mean value of the dataset metric evaluated.

If a model is robust in its predictions, and the dataset does not involve excessive error measurements, most of the data records should be located on a William's diagram within the following limits:

$$0 \leq H \leq 0.1725$$
$$-3 \leq SR \leq 3$$

It is obvious in Figs. 12 and 13 that most of the data points (~94%–95%) for both ω and λ fall within the feasible regions.

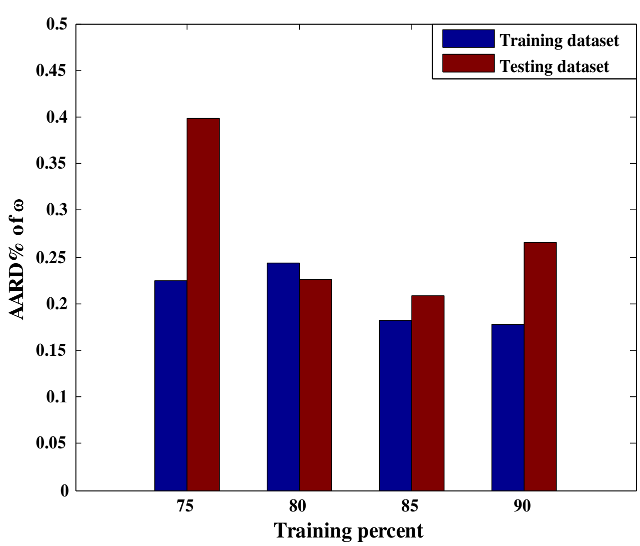

Fig. 7 Sensitivity of the MLS-SVR model in prediction of ω to the distribution of training and testing subsets

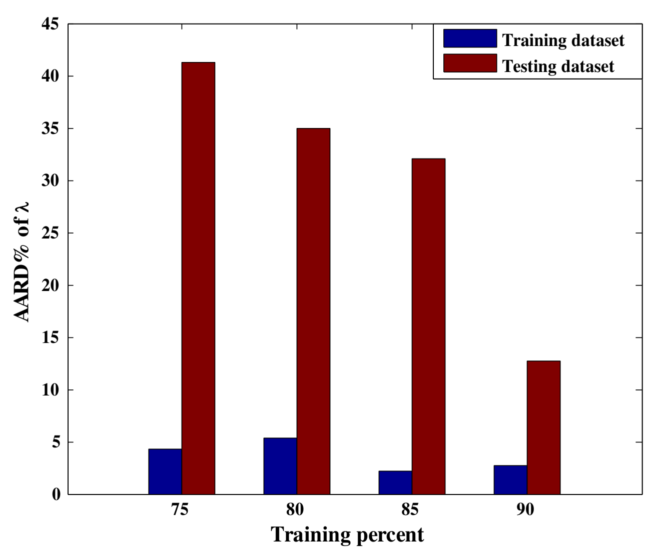

Fig. 8 Sensitivity of the MLS-SVR model in prediction of λ to the distribution of training and testing subsets

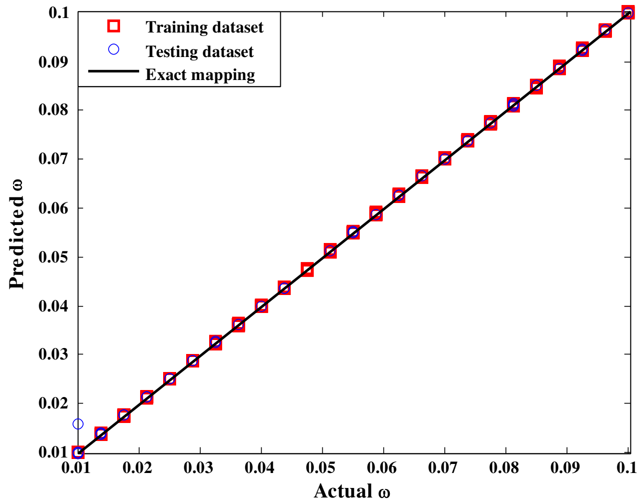

Fig. 9 Comparison of actual values of ω versus predicted values by the MLS-SVR model

There are only 28 data records identified as possible outliers for ω, and 32 data records identified as possible outliers for λ. This observation indicates high levels of statistical confidence in the predictions generated by the MLS-SVR model.

Performance of the developed MLS-SVR model applied to a published signal

Figure 14 presents a pressure derivative graph for a drawdown test performed on a horizontal well in a dual-porosity reservoir (Cheng 2004). The values of storativity ratio and interporosity flow coefficient for this signal are reported in that work as 0.1 and 1×10^{-6}, respectively.

The developed MLS-SVR approach is applied to this signal. and predicts storativity ratio and interporosity flow

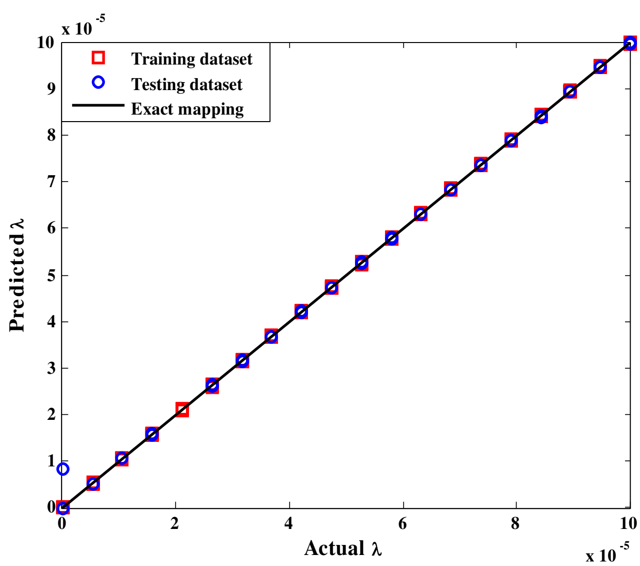

Fig. 10 Comparison of actual values of λ versus predicted values by the MLS-SVR model

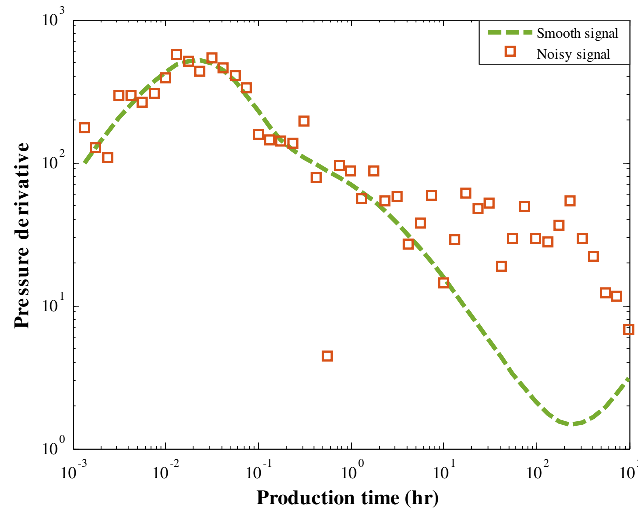

Fig. 11 Pressure derivative curve for a noisy record data from the dual-porosity NFR

coefficient as 0.08419 and 1.16×10^{-6}. The MLS-SVR model is clearly able to identify and predict that signal accurately.

Conclusion

In this study, a multi-output least squares support vector regression model developed and successfully applied to predict two key parameters of naturally fractured reservoirs, i.e., interporosity flow coefficient and storativity ratio. A huge databank having 500 PTRs is generated using an analytical solution of partial differential equations defining fluid flow in a fractured reservoir. These PRT then converted to the pressure derivative curves to facilitate the MLS-SVR for discrimination among different values of ω and λ. AARD of 0.177% and 2.755% were obtained for trained ω and λ predictions by the MLS-SVR, respectively. The leverage method identified that up to about 6% of the predictions made by the method were potential outliers. These results suggest that the MLS-SVR method developed and applied in this study to a dataset of pressure derivative curves provides a useful alternative to conventional PTA methods, mainly when applied to more complex reservoirs dual-porosity natural fractured reservoir flow system evaluated.

Appendix 1. Multi-output least squares support vector regression

Support vector regression maps nonlinear patterns into higher dimensional feature space that can approach infinite dimensions (Khandelwal and Kankar 2011; Tikhamarine et al. 2019; Quan et al. 2020). It then applies linear regression to the mapped feature space (Chao et al. 2018; Xu and Chen

Table 3 The prediction performance of the MLS-SVR model applied to example noisy data records of well-test pressure derivative curves for a dual-porosity NFR

Real value of ω	Predicted by MLS-SVR	Real value of λ	Predicted by MLS-SVR
0.0100	0.0101	1×10^{-8}	1.09×10^{-8}
0.0138	0.013	5.273×10^{-6}	5.32×10^{-6}
0.0175	0.018	1.053×10^{-5}	1.13×10^{-5}
0.0213	0.021	1.580×10^{-5}	1.6×10^{-5}
0.0250	0.024	2.106×10^{-5}	2×10^{V5}
0.0288	0.028	2.632×10^{-5}	2.32×10^{-5}
0.0325	0.033	3.159×10^{-5}	3.01×10^{-5}
0.0363	0.036	3.685×10^{-5}	3.89×10^{-5}
0.0400	0.038	4.211×10^{-5}	4.01×10^{-5}
0.0438	0.042	4.737×10^{-5}	4.35×10^{-5}

2019). Considering a dataset of N records existing in multi-dimensional feature space, its ith data record or element can be expressed as:

$$[(x_i, y_i), i = 1, 2, \ldots, N]$$

Here, x_i and y_i represent an actual and predicted value of the ith data record. For such a data set, the support vector regression can be expressed as follows (Xu et al. 2013):

$$F(x) =< w, x > +b \tag{9}$$

where <,> denotes the dot product of the matrix elements involving all the x data records, w is the weight vector of the SVR regression function, and b is an intercept of the SVR regression function.

Accepting a certain level of error (ε), the objective of SVR is to establish a function $F(x)$ that estimates the values of y from x data for a training dataset that maintains deviations at or less than the value specified for ε. The risk function (R) can then be solved using appropriate optimization techniques (Chen et al. 2017; Deng et al. 2019; Liu et al. 2019a; Cao et al. 2020a, b, c; Qu et al. 2020).

$$R(f) = \frac{1}{N}\sum_{i=1}^{N} L(f(x_i)-y_i) + \frac{1}{2}\|w\|^2 \tag{10}$$

where:

$$L(f(x)-y) = \begin{cases} \|L(f(x)-y)\|-\varepsilon & \text{if } |f(x)-y|\rangle 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Equation (4) expresses an error insensitive loss function. ε determines the regression's precision by essentially defining the radius of a cylinder surrounding the regression function, $f(x)$, within which acceptable values may exist. By substitution of Eq. (3) into Eq. (2), it is possible to determine functions that fit the records of the SVR training subset with deviations of no more than ε.

The acceptable error components associated with SVR minimization can be further defined as follows (Smola and Schölkopf 2004; Chen et al. 2019):

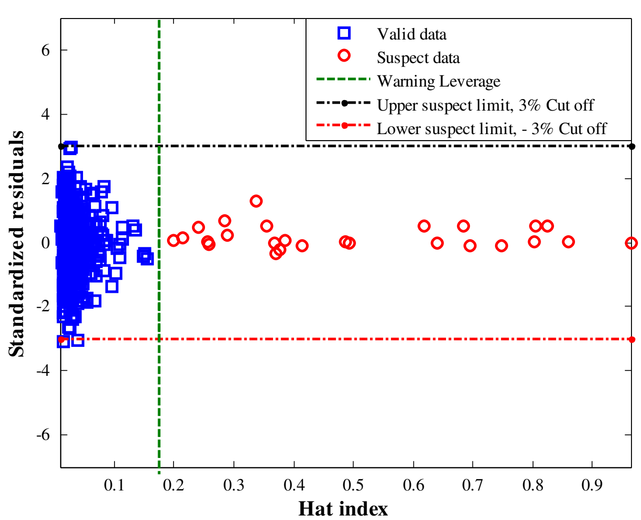

Fig. 12 William's plot for detection of suspected outlier predictions of ω by the MLS-SVR model. The dashed vertical line represents the H^* value calculated using Eq. (8)

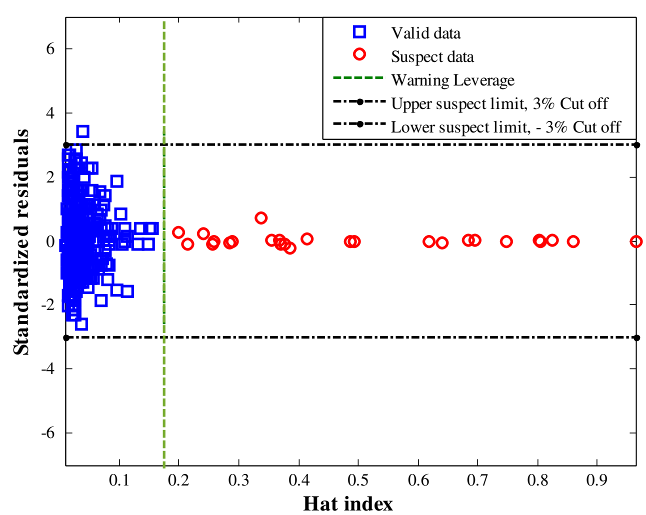

Fig. 13 William's plot for detection of suspected outlier predictions of λ by the MLS-SVR model for the dual-porosity NFR dataset analyzed. The dashed vertical line represents the H^* value calculated using Eq. (8)

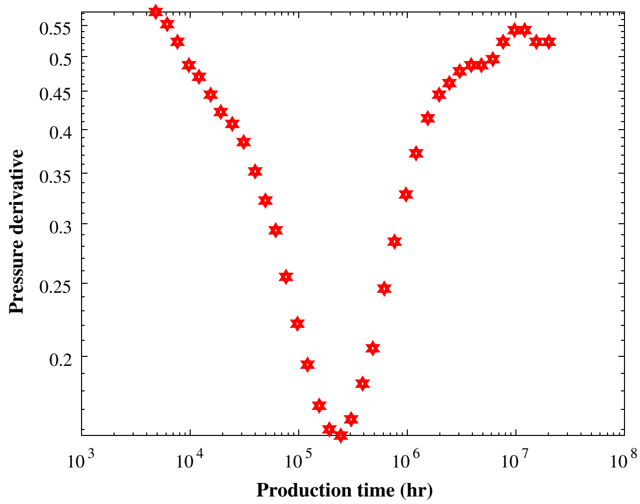

Fig. 14 Pressure derivative graph from literature (Cheng 2004)

$$\text{Minimize}\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi + \xi^*) \tag{12}$$

This minimization involves the constraints listed in Eq. (5):

$$\text{s.t.}\begin{cases} y_i - \langle w, x_i\rangle + b < \varepsilon + \xi_i \\ \langle w, x_i\rangle + b - y_i < \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0 \\ i = 1,2,\ldots,N \end{cases} \tag{13}$$

where C is a positive regularization constant, ξ_i and ξ_i^* are positive slack variables.

C is a metric that establishes a trade-off between a solution's ability to be generalized across all the elements of a data subset (e.g., the SVR training subset) versus achieving acceptable levels of accuracy (expressed in terms of error tolerance by ε). ξ_i and ξ_i^* quantify the distance from the values to theboundaries values of the error cylinder defined by ε.

SVR can then be expressed as a dual problem in the form of Eq. (6) to be maximized (Burges 1998; Bian et al. 2016):

$$\text{maxmize}\ \frac{1}{2} \times \sum_{i,j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i x_j) - \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} y_i(\alpha_i + \alpha_i^*) \tag{14}$$

$$\text{s.t.}\begin{cases} \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* > 0 \\ i = 1, 2, \ldots, N \end{cases} \tag{15}$$

Here, α_i and α_i^* are Lagrange multipliers derived from quadratic equation solutions.

The SVR single space function can then be mathematically expressed by Eq. (8):

$$f(x) = (\alpha_i - \alpha_i^*)\langle x_i, x\rangle + b \tag{16}$$

The SVR dual space regression function is expressed by Eq. (9):

$$f(x) = (\alpha_i - \alpha_i^*)k(x_i, x) + b \tag{17}$$

where $k(x_i, x)$ indicates the kernel function that satisfies Mercer's conditions.

Those data records in the dataset determined by Eq. (9) to have non-zero coefficients are the support vectors. A kernel function commonly applied in Eq. (9) is the radial basis function (RBF) expressed as Eq. (10):

$$k(x_i, x_j) = \exp(-\gamma\|x_i, x_j\|), (\gamma > 0) \tag{18}$$

where γ shows the width of the RBF.

The critical variables in establishing acceptable SVR optimization solutions (regression functions) and the complexity of those solutions are ε (Eq. 3), C (Eq. 4), and γ (Eq. 10). SVR optimization, therefore, focuses upon optimizing the variables ε, C, and γ.

In MLS-SVR, each data record in the dataset has multiple independent-dependent variables (Yan et al. 2020). For the case of our study, the digitized pressure derivative curve with 28 points is the independent variable, while ω and λ constitute the target vector.

Nomenclature *V,* bulk volume; *c,* compressibility; *k,* permeability; *L,* half-length of the horizontal well; *p,* pressure; *t,* time; *R,* risk function; *w,* weight vector; *b,* intercept of the SVR regression function; $k(x_i, x)$, kernel function; *N,* number of data records; *H,* Hat matrix; *X,* matrix of standardized residuals; *H*,* warning leverage value; *n,* number of model parameters; *m,* number of data records

Greek symbols λ, interporosity flow coefficient; ω, storativity ratio; ϕ, porosity; α, shape factor; Δ, difference; ε, error; ξ_i, slack variable; α_i, Lagrange multiplier; γ, width of the RBF

Abbreviations *NFR,* naturally fractured reservoirs; *PTR,* pressure transient response; *MLS-SVR,* multi-output least squares support vector regression; *PD,* pressure derivative; *AARD%,* absolute average relative deviations; R^2, regression coefficients; *PTA,* pressure transient analyses; *ANN,* artificial neural networks; *RBF,* radial basis function; *MSE,* mean square error; *PV,* parameter value; *SR,* cutoff value for standardized residuals; *hr,* hour

Superscripts/subscripts *f,* fracture; *m,* matrix; *t,* total; *act.,* actual value; *cal.,* calculated value; *T,* transpose; -1, Inverse; *D,* dimensionless

Declarations

Conflict of interest The author (s) declare that they have no competing interests.

References

Abdel-Fattah M, Gameel M, Awad S, Ismaila A (2015) Seismic interpretation of the aptian alamein dolomite in the razzak oil field, western desert, Egypt. Arab J Geosci 8:4669–4684

Alajmi MN, Ertekin T (2007) The development of an artificial neural network as a pressure transient analysis tool for applications in double-porosity reservoirs. In: Asia Pacific oil and gas conference and exhibition. Society of Petroleum Engineers, 30th October–1 November, Jakarta, Indonesia

Barenblatt GI, Zheltov YP (1960) Fundamental equations for the flow of homogeneous fluids through fissured rocks. In: Doklady Akademii Nauk. Russ Acad Sci 132:545–548

Bian XQ, Han B, Du ZM, Jaubert JN, Li MJ (2016) Integrating support vector regression with genetic algorithm for CO2-oil minimum miscibility pressure (MMP) in pure and impure CO2 streams. Fuel 182: 550–557

Bourdet D (2002) Well test analysis: the use of advanced interpretation models. Elsevier, Amsterdam

Bourdet D, Ayoub JA, Pirard YM (1989) Use of pressure derivative in well test interpretation. SPE Form Eval 4:293–302

Burges CJC (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discovery 2:121–167

Cao B, Dong W, Lv Z, Gu Y, Singh S, Kumar P (2020a) Hybrid microgrid many-objective sizing optimization with fuzzy decision. IEEE Trans Fuzzy Syst 28:2702–2710

Cao B, Zhao J, Gu Y, Ling Y, Ma X (2020b) Applying graph-based differential grouping for multiobjective large-scale optimization. Swarm Evol Comput 53:100626

Cao Y, Li Y, Zhang G, Jermsittiparsert K, Nasseri M (2020c) An efficient terminal voltage control for PEMFC based on an improved version of whale optimization algorithm. Energy Rep 6:530–542

Chao L, Zhang K, Li Z, Zhu Y, Wang J, Yu Z (2018) Geographically weighted regression based methods for merging satellite and gauge precipitation. J Hydrol 558:275–289

Chen Y, He L, Guan Y, Lu H, Li J (2017) Life cycle assessment of greenhouse gas emissions and water-energy optimization for shale gas supply chain planning based on multi-level approach: case study in Barnett, Marcellus, Fayetteville, and Haynesville shales. Energy Convers Manag 134:382–398

Chen H, Qiao H, Xu L, Feng Q, Cai K (2019) A fuzzy optimization strategy for the implementation of RBF LSSVR model in vis–NIR analysis of Pomelo maturity. IEEE Trans Ind Inform 15:5971–5979

Chen H, Chen A, Xu L, Xie H, Qiao H, Lin Q, Cai K (2020a) A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. Agric Water Manag 240:106303

Chen H, Zhang G, Fan D et al (2020b) Nonlinear lamb wave analysis for microdefect identification in mechanical structural health assessment. Measurement 108026

Chen Y, Li J, Lu H, Yan P (2021) Coupling system dynamics analysis and risk aversion programming for optimizing the mixed noise-driven shale gas-water supply chains. J Clean Prod 278:123209

Cheng Y (2004) Pressure transient testing and productivity analysis for horizontal wells, (Doctoral dissertation, Texas A&M University).

Cheng X, He L, Lu H, Chen Y, Ren L (2016) Optimal water resources management and system benefit for the Marcellus shale-gas reservoir in Pennsylvania and West Virginia. J Hydrol 540:412–422

Ciftci E (2018) A practical approach for the interpretation of flowing well tests. Arab J Geosci 11:761

Daolun LI, Xuliang LIU, Wenshu ZHA et al (2020) Automatic well test interpretation based on convolutional neural network for a radial composite reservoir. Pet Explor Dev 47:623–631

Deng Y, Zhang T, Sharma BK, Nie H (2019) Optimization and mechanism studies on cell disruption and phosphorus recovery from microalgae with magnesium modified hydrochar in assisted hydrothermal system. Sci Total Environ 646:1140–1154

Egya D, Geiger S, Corbett P, March R (2017) The effect of fracture skin, network connectivity, and network size on well-test responses in naturally fractured reservoirs. In: 79th EAGE Conference and Exhibition 2017. European Association of Geoscientists & Engineers, 12-15 June, Paris, France

Engler T, Tiab D (1996) Analysis of pressure and pressure derivative without type curve matching, 4. Naturally fractured reservoirs. J Pet Sci Eng 15:127–138

Escobar FH, Zhao Y-L, Urazán C, Trujillo CM (2018) Pressure and pressure derivative interpretation for horizontal wells in compressible formations. J Geophys Eng 15:1551–1560

Eslamloueyan R, Vaferi B, Ayatollahi S (2010) Fracture characterizations from well testing data using artificial neural networks. In: 72nd European Association of Geoscientists and Engineers Conference and Exhibition 2010: A New Spring for Geoscience. Incorporating SPE EUROPEC 2010, 14-17 June, Barcelona, Spain

Ghaffarian N, Eslamloueyan R, Vaferi B (2014) Model identification for gas condensate reservoirs by using ANN method based on well test data. J Pet Sci Eng:123

Gramatica P (2007) Principles of QSAR models validation: internal and external. QSAR Comb Sci 26:694–701

Guo H, Li X, Zhu Q, Zhang Z, Liu Y, Li Z, Wen H, Li Y, Tang J, Liu J (2020a) Imaging nano-defects of metal waveguides using the microwave cavity interference enhancement method. Nanotechnology 31: 455203

Guo T, Tang S, Liu S et al (2020b) Physical Simulation of Hydraulic Fracturing of Large-Sized Tight Sandstone Outcrops. SPE J

He L, Chen Y, Zhao H, Tian P, Xue Y, Chen L (2018) Game-based analysis of energy-water nexus for identifying environmental impacts during shale gas operations under stochastic input. Sci Total Environ 627:1585–1601

Huang S, Yao Y, Zhang S, Ji J, Ma R (2018) Pressure transient analysis of multi-fractured horizontal wells in tight oil reservoirs with consideration of stress sensitivity. Arab J Geosci 11:285

Khan M, Arif M, Ali N, Yaseen M, Ahmed A, Siyar SM (2016) Petrophysical parameters and modelling of the Eocene reservoirs in the Qadirpur area, Central Indus Basin, Pakistan: implications from well log analysis. Arab J Geosci 9:425

Khandelwal M, Kankar PK (2011) Prediction of blast-induced air overpressure using support vector machine. Arab J Geosci 4:427–433

Landa JL, Horne RN, Kamal MM, Jenkins CD (2000) Reservoir characterization constrained to well-test data: a field example. SPE Reserv Eval Eng 3:325–334

Li T, Xu M, Zhu C, Yang R, Wang Z, Guan Z (2019) A deep learning approach for multi-frame in-loop filter of HEVC. IEEE Trans Image Process 28:5663–5678

Li Z, Liu H, Dun Z, Ren L, Fang J (2020) Grouting effect on rock fracture using shear and seepage assessment. Constr Build Mater 242: 118131

Liu E, Lv L, Yi Y, Xie P (2019a) Research on the steady operation optimization model of natural gas pipeline considering the combined operation of air coolers and compressors. IEEE Access 7: 83251–83265

Liu Y, Chen S, Guan B, Xu P (2019b) Layout optimization of large-scale oil–gas gathering system based on combined optimization strategy. Neurocomputing 332:159–183

Liu X, Li D, Yang J, Zha W, Zhou Z, Gao L, Han J (2020) Automatic well test interpretation based on convolutional neural network for infinite reservoir. J Pet Sci Eng 195:107618

Lv Z, Qiao L (2020) Deep belief network and linear perceptron based cognitive computing for collaborative robots. Appl Soft Comput 92: 106300

Moosavi SR, Qajar J, Riazi M (2018a) A comparison of methods for denoising of well test pressure data. J Pet Explor Prod Technol 8: 1519–1534

Moosavi SR, Vaferi B, Wood DA (2018b) Applying orthogonal collocation for rapid and reliable solutions of transient flow in naturally fractured reservoirs. J Pet Sci Eng 162:166–179

Moosavi SR, Vaferi B, Wood DA (2020) Auto-detection interpretation model for horizontal oil wells using pressure transient responses. Adv Geo-Energy Res 4:305–316

Executegh M, Vaferi B, Riazi M (2019) Orthogonal collocation method for solving the diffusivity equation: Application on dual porosity reservoirs with constant pressure outer boundary. ASME J Energy Resour Technol 141:042001

Qian J, Feng S, Li Y, Tao T, Han J, Chen Q, Zuo C (2020a) Single-shot absolute 3D shape measurement with deep-learning-based color fringe projection profilometry. Opt Lett 45:1842–1845

Qian J, Feng S, Tao T, Hu Y, Li Y, Chen Q, Zuo C (2020b) Deep-learning-enabled geometric constraints and phase unwrapping for single-shot absolute 3D shape measurement. APL Photonics 5: 46105

Qu S, Han Y, Wu Z, Raza H (2020) Consensus modeling with asymmetric cost based on data-driven robust optimization. Gr Decis Negot 19:1–38

Quan Q, Hao Z, Xifeng H, Jingchun L (2020) Research on water temperature prediction based on improved support vector regression. Neural Comput Applic 28:1–10

Rousseeuw PJ, Leroy AM (2005) Robust regression and outlier detection. John wiley & sons, New York

Şahin AU, Çiftçi E (2016) An integration-based estimation approach for the determination of slug test parameters under various flow geometries. Arab J Geosci 9:639

Saidi AM (1983) Simulation of naturally fractured reservoirs. SPE Res Eng 3(02):638–648

Shi K, Wang J, Zhong S, Tang Y, Cheng J (2020) Non-fragile memory filtering of TS fuzzy delayed neural networks based on switched fuzzy sampled-data control. Fuzzy Sets Syst 394:40–64

Sierra O (1986) Fundamentals of well-log interpretation: The Interpretation of Logging Data, Elsevier, Amsterdam

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Stat Comput 14:199–222

Su Z, Liu E, Xu Y, Xie P, Shang C, Zhu Q (2019) Flow field and noise characteristics of manifold in natural gas transportation station. Oil Gas Sci Technol d'IFP Energies Nouv 74:70

Tiab D (1989) Direct type-curve synthesis of pressure transient tests. In: Low Permeability Reservoirs Symposium. Society of Petroleum Engineers

Tiab D (1994) Analysis of pressure and pressure derivative without type-curve matching: Vertically fractured wells in closed systems. J Pet Sci Eng 11:323–333

Tiab D, Donaldson EC (2015) Petrophysics: theory and practice of measuring reservoir rock and fluid transport properties, 3rd ed. Gulf professional publishing, Amsterdam

Tikhamarine Y, Souag-Gamane D, Kisi O (2019) A new intelligent method for monthly streamflow prediction: hybrid wavelet support vector regression based on grey wolf optimizer (WSVR–GWO). Arab J Geosci 12:540

Vaferi B, Eslamloueyan R, Ayatollahi S (2011) Automatic recognition of oil reservoir models from well testing data by using multi-layer perceptron networks. J Pet Sci Eng 77:254–262

Vaferi B, Eslamloueyan R, Ghaffarian N (2016) Hydrocarbon reservoir model detection from pressure transient data using coupled artificial neural network-Wavelet transform approach. Appl Soft Comput J 47:63–75

Wang F (2016) Continuous reservoir model calibration with time-dependent reservoir properties diagnosed from long-term down-hole transient pressure data. Arab J Geosci 9:254

Warren JE, Root PJ (1963) The behavior of naturally fractured reservoirs. Soc Pet Eng J 3:245–255

Xu X, Chen L (2019) Projection of long-term care costs in China, 2020–2050: based on the Bayesian quantile regression method. Sustainability 11:3530

Xu S, An X, Qiao X, Zhu L, Li L (2013) Multi-output least-squares support vector regression machines. Pattern Recogn Lett 34:1078–1084

Xu M, Li T, Wang Z, Deng X, Yang R, Guan Z (2018) Reducing complexity of HEVC: a deep learning approach. IEEE Trans Image Process 27:5044–5059

Yan J, Pu W, Zhou S, Liu H, Greco MS (2020) Optimal resource allocation for asynchronous multiple targets tracking in heterogeneous radar networks. IEEE Trans Signal Process 68:4055–4068

Yang S, Deng B, Wang J et al (2019) Scalable digital neuromorphic architecture for large-scale biophysically meaningful neural network with multi-compartment neurons. IEEE Trans Neural Netw Learn Syst 31:148–162

Yang Y, Liu J, Yao J, Kou J, Li Z, Wu T, Zhang K, Zhang L, Sun H (2020) Adsorption behaviors of shale oil in kerogen slit by molecular simulation. Chem Eng J 387:124054

Yue H, Wang H, Chen H et al (2020) Automatic detection of feather defects using Lie group and fuzzy Fisher criterion for shuttlecock production. Mech Syst Signal Process 141:106690

Zhang X, Wang Y, Chen X et al (2018) Decentralized adaptive neural approximated inverse control for a class of large-scale nonlinear hysteretic systems with time delays. IEEE Trans Syst Man, Cybern Syst 49:2424–2437