PRODUCTION DATA ANALYSIS AND PRODUCTION FORECASTING FOR UNCONVENTIONAL RESERVOIRS USING MACHINE LEARNING ALGORITHMS

A Dissertation

by

YUEWEI PAN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	John Lee
Committee Members,	Marcelo Laprea-Bigott
	Peter Bastian
	Theodora Chaspari
Head of Department,	Jeff Spath

December 2019

Major Subject: Petroleum Engineering

ABSTRACT

Production data analysis for low permeability shale reservoirs is crucial in characterizing flow regimes as well as reservoir properties, and the forecasting of production is essential for portfolio and reservoir management. However, traditional methods have failed due to incorrect physics or complicated convolution from the well control history. In this research, we investigate several novel machine learning approaches and their variations to better facilitate production surveillance as well as accelerate the analysis of variable and noisy historical production data.

Current production data analyses are on the basis of either simple constant drawdown-pressure rate solutions or constant rate pressure-drop solutions in which we cannot handle field data variation. Therefore, deconvolution methods have been developed to transform the variable pressure variable rate profiles into those two simple solutions. Pressure-rate or rate-pressure deconvolution is an ill-posed, complex time-series problems. Much research has indicated that Echo-State Networks (ESN) and Long Short-Term Memory (LSTM) are useful for dynamic time-series problems, however, the connections between physics and machine-learning-based solutions remain unpublished. Thus, one of the motivations of this research is to establish the connection between transient physics and machine learning algorithms specifically using ESN and LSTM paradigms.

Traditional decline curve analysis (DCA) models have played an important role in the oil and gas industry for nearly a century because of their simplicity. However, forecasting models with too much simplicity sometimes rely on unnecessary assumptions and fail to honor important realities. Unlike those DCA methods which rely solely on rate histories, this research first introduces novel analytic methods to accomplish forecasting using pressure-rate-time information while still maintaining the reduced complexities. By switching the application from rate transient analysis to pressure transient analysis, these analytic models are demonstrated to be useful in identifying the reservoir model with only pressure-rate information.

In general, the purpose of this study is to present alternative methodologies for deconvolution to facilitate the production data analysis. We forecast future production in addition to applying the ‘inverse’ process for reservoir characterization. We also utilize the merits of physics-based training features to further enhance the production surveillance and finally, we illustrate the proposed workflows for production analysis, production forecasting and production surveillance using variants under the ESN and LSTM paradigms with synthetic cases and field examples.

DEDICATION

To my parents and my fiancée who have always supported and loved me unconditionally.

ACKNOWLEDGEMENTS

First of all, I am grateful to my advisor and committee chair, Dr. John Lee for guiding me throughout the entire four and a half years of study. I feel fortunate to have such a challenging and forward-looking research topic to enrich my research career. Dr. Lee provided extraordinary support when I faced “dilemmas”, I truly express my whole-hearted gratitude to Dr. Lee by quoting “O My Captain! My Captain!” by Walt Whitman.

Secondly, I also thank Prof. Peter Bastian and Dr. Marcelo Laprea-Bigott for their industry expertise and full support to serve as the members of my committee. Additionally, I would like to extend my gratitude to Dr. Theodora Chaspari for sharing her expertise in machine learning for signal processing.

Furthermore, many thanks to my graduate colleagues, the faculty and the staff in the Petroleum Engineering Department for their contributions to enhance my research progress and my student life experiences. I especially appreciate the contributors of the University Land database which fundamentally accelerated my research progress and made my entire project unique.

Finally, I am grateful to my family for their encouragements and endless support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor John Lee, Professor Peter Bastian, Professor Marcelo Laprea-Bigott of the Department of Petroleum Engineering and Professor Theodora Chaspari of the Department of Computer Science and Engineering.

The data analysis in Chapter IV was provided by Professor Marcelo Laprea-Bigott using University Land database.

I conducted all other work for the dissertation independently.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

1.1 Background Overview

Within recent decades, horizontal wells with multistage hydraulic fractures as the major driving force for economic growth have been ubiquitously applied in unconventional reservoirs. Production analysis based on historical rate/pressure data is a basic process to determine reservoir properties and identify the reservoir model either depending on Pressure Transient Analysis (PTA) or Rate Transient Analysis (RTA). The outcome from RTA could be further used in forecasting future production. The theoretical behavior of unconventional reservoirs that produce economically with massive fracture treatments and complex fracture networks is shown in Fig. 1 and Fig. 2. The rapid rate change induced by unstable operating conditions as well as the long-term transient dominated behavior complicate production analysis and forecasting for unconventional reservoirs. With the emerging digitized technology and intelligent algorithms, production surveillance and monitoring have been raised as fundamental pillars for production data analysis, reservoir diagnostics and company-wise operational costs checking as illustrated.

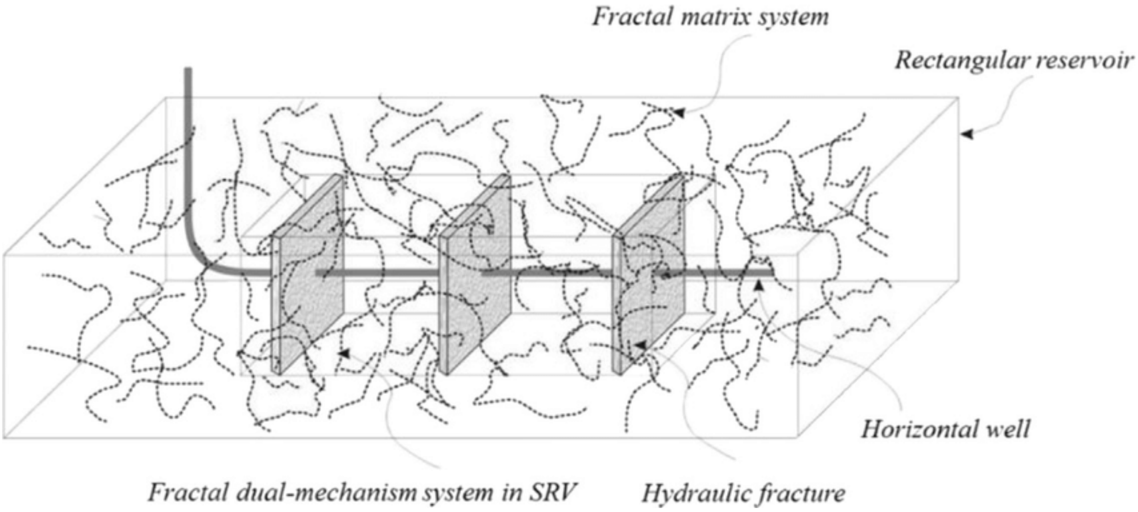


Fig. 1-A theoretical model of a multiple transverse fractured horizontal well in a rectangular reservoir with complex fracture networks (Wang et al. 2015).

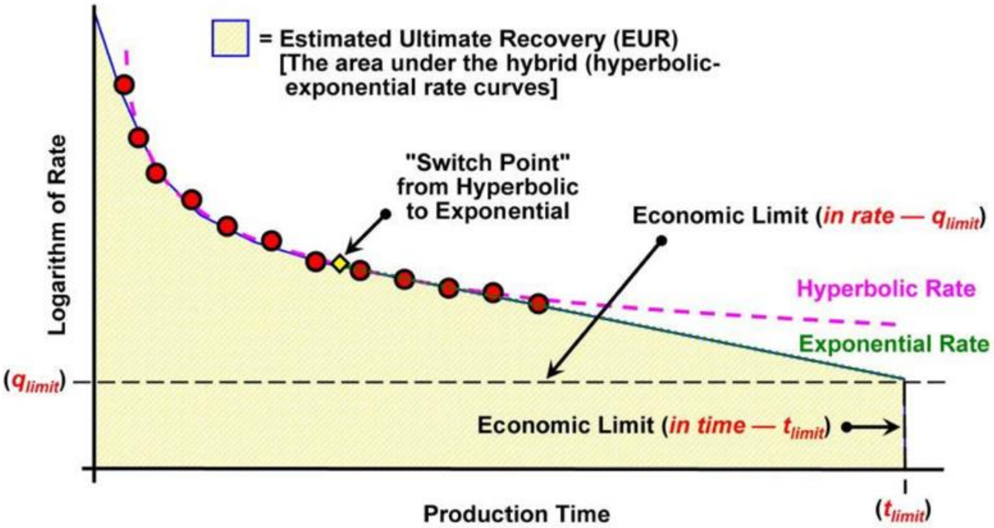


Fig. 2-A theoretical approach for unconventional reservoir production forecasting and prediction of estimate ultimate recoveries (EUR) (DeGolyer and MacNaughton, 2019).

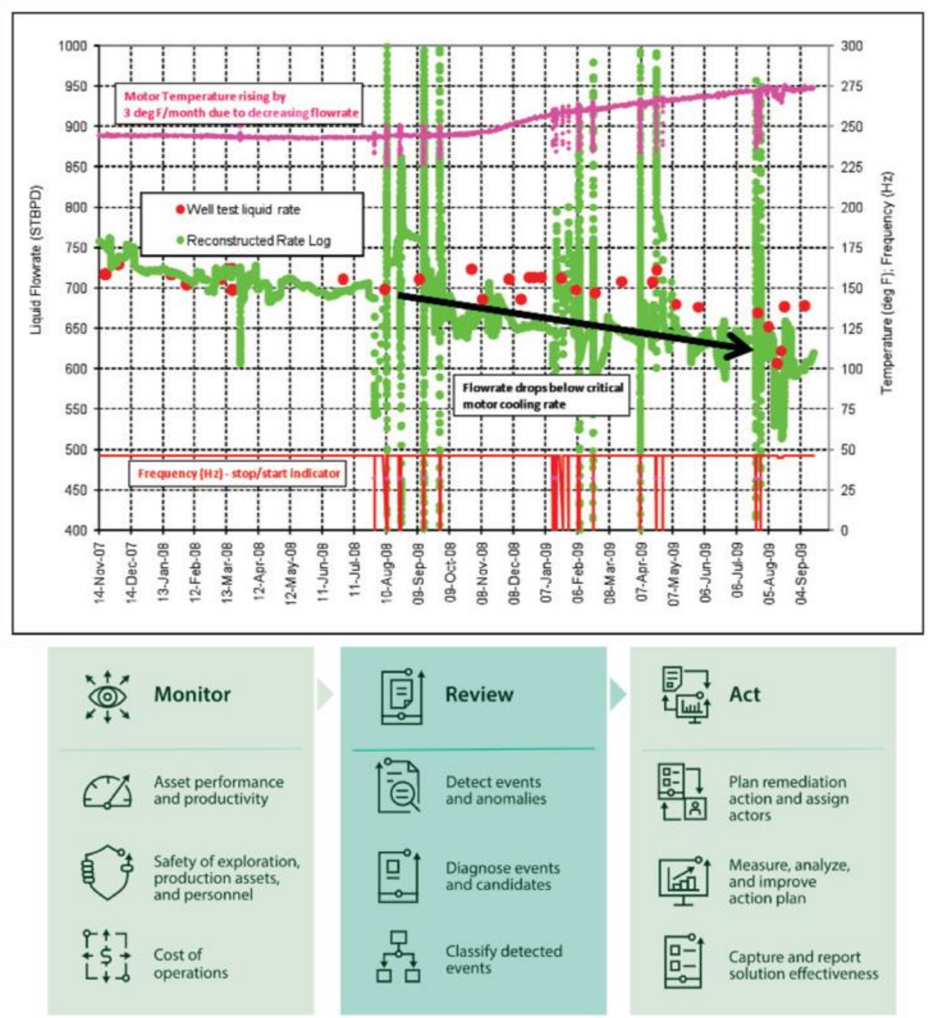


Fig. 3-An example of downhole production monitoring followed by some real-time analysis (Camilleri, 2010) and a schematic workflow of data analyses, reservoir diagnostics and strategy making (Infosys, Automate Workflow to Meet Production Targets).

Common techniques for production data analysis are either to establish a numerical model to match history or use analytical models to recognize flow patterns and analytically estimate reservoir properties. Numerical history matching, a ‘feed-forward’ process shown in Fig. 4 (which requires prior knowledge of geological information), is widely applied in modern production data analysis due to its accuracy and versatility. Once the numerical model has been validated, future production can be simulated with the forecasting procedure. In contrast, the analytical method such as using diagnostic plots, is an inverse process for reservoir parameters estimation as indicated in Fig. 5. Then, traditional decline curve analysis (DCA) models illustrated in Fig. 6 such as Arp's decline model (Arps 1945), the stretched exponential decline model (Valko and Lee 2010) could be further utilized for forecasting. Production data gathering from downhole relies on real-time monitoring equipped with permanent downhole gauges (PDG). It enables engineers to diagnose the production status and rapidly respond to the changed the conditions (Mittal 2018). Driven by the reduction of data storage expenses and the development of machine learning algorithms, downhole real-time data is no longer limited to decision making, instead, a more -collaborative, -automated and -efficient process can be implemented without increasing downtime (Mittal 2018). Real-time pressure-rate information inevitably carries missing values, aberrant segments, outliers, and systematic noise due to unexpected failures, routinely system offline as well as random errors (Cao et al. 2018) which can be clearly depicted in Fig. 7. In section 2.1, we propose novel analytic methods to analyze production data using both PTA and RTA, and to forecast production using reduced-complexity models as the basis in RTA while achieving results comparable with full field modeling and inversion. In section 2.2, we propose a deep learning workflow to reconstruct pressure-rate-time information from production surveillance. It captures the attributes of the flow mechanisms by considering spatial dependency, inherits physics-based training features from section 2.1, further enables engineers to use entire production histories even with outliers, aberrant segments, and augmented noise, and thus improves production data analysis and forecasting for unconventional reservoirs.

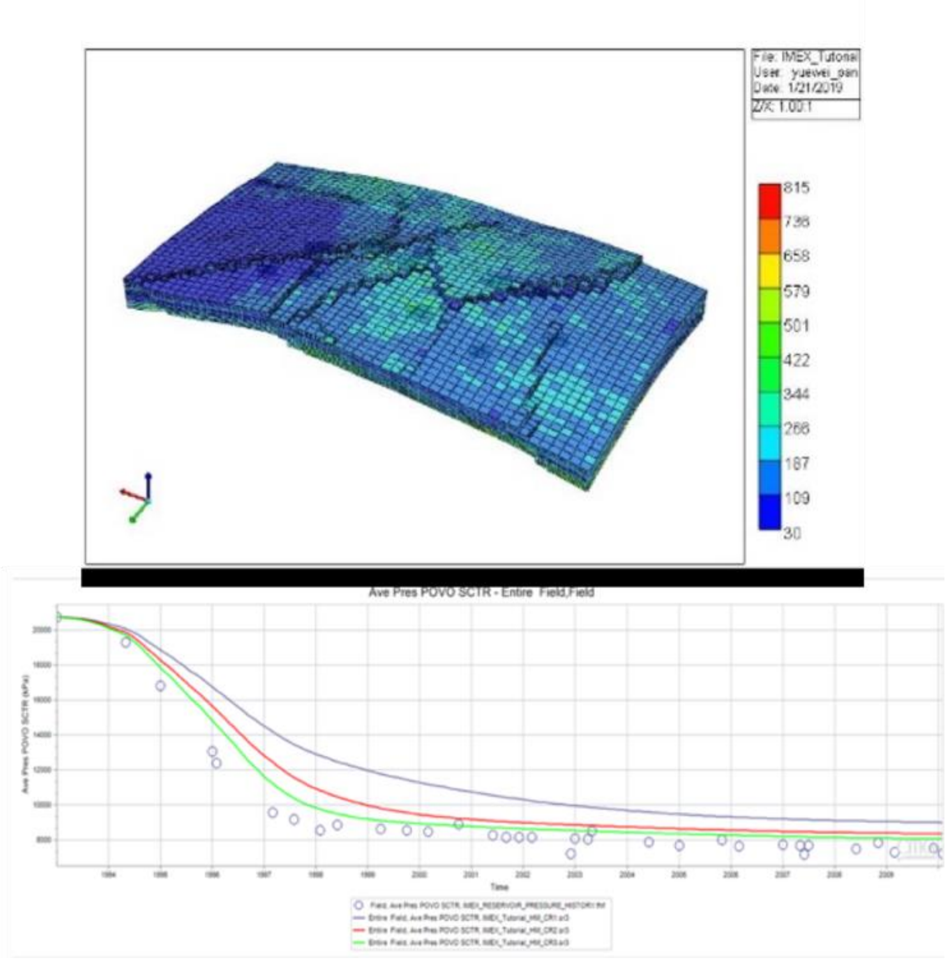


Fig. 4-Numerical simulation for history matching using CMG IMEX.

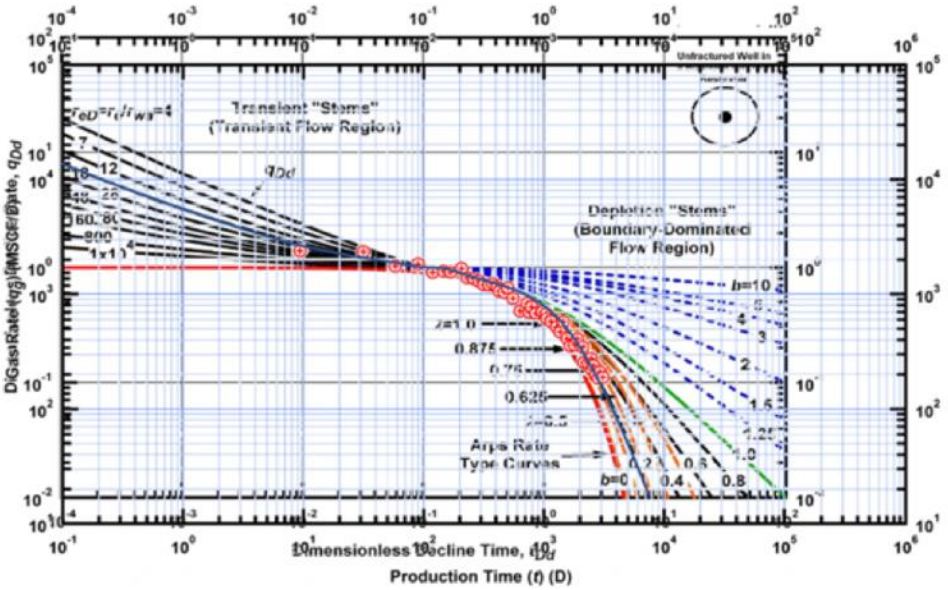


Fig. 5-A practical analytical process using Fetkovich type curve to match production data.

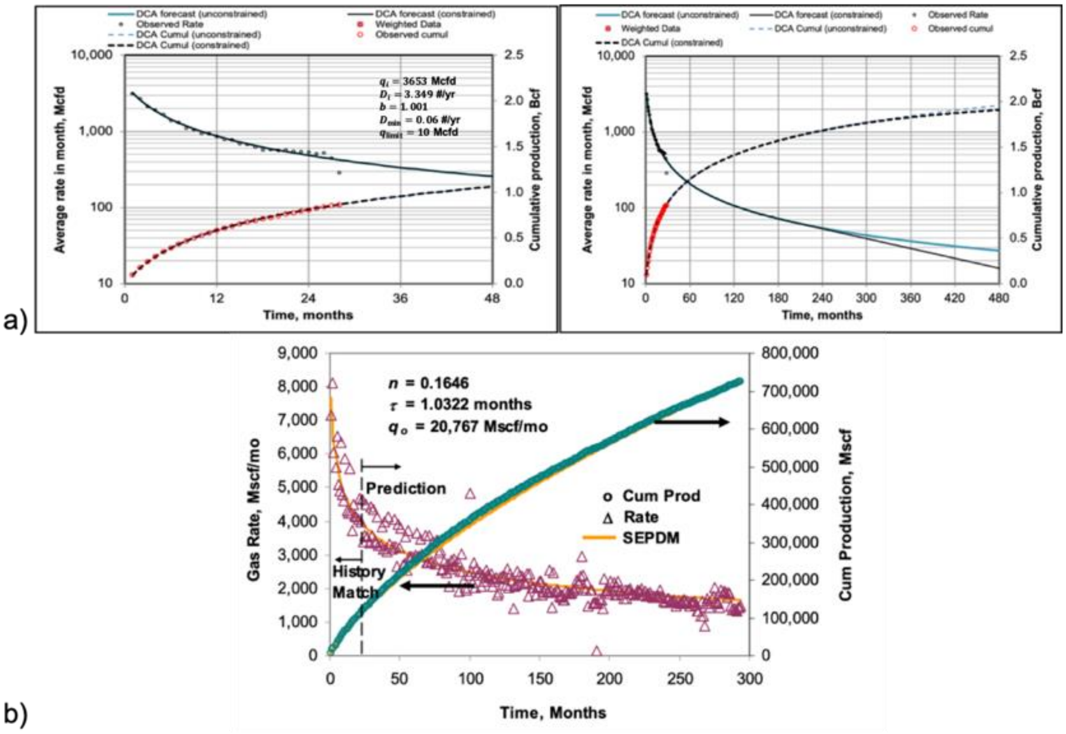


Fig. 6 (a)-Arps decline fit for the Canadian unconventional well; (b)-Stretched exponential decline model fit for the Marcellus well at the bottom (Kabir et al. 2011).

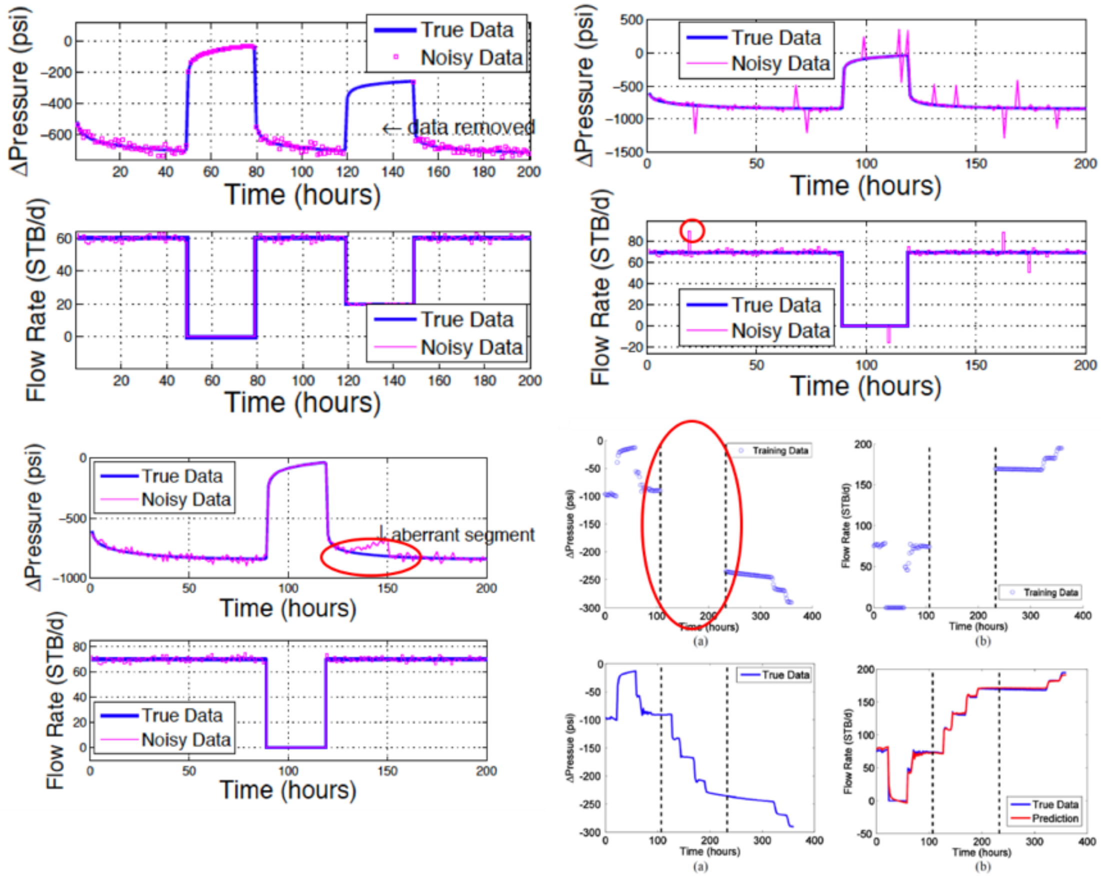


Fig. 7-Permanent downhole gauge sampling noisy data with outliers, aberrant segment and outliers (Tian and Horne 2015).

1.2 Motivation and Objectives

Applying analytical solutions for production data analysis (PDA), with the help of pressure transient analysis (PTA) (Lee et al. 2003) or rate transient analysis (RTA) is usually an ill-conditioned ‘inverse’ process that is constrained in application due to variations in pressure or rate and the effect of backpressure in the wellbore (Kuchuk et al. 2005). To achieve the mathematically preferable constant pressure or constant rate solution, deconvolution has been proposed and applied to pressure/rate data. An issue arising in the traditional deconvolution in PTA is addressed by Liu and Horne (Liu et al.

2013). Systematic curve-fitting techniques were required because of the additional impact of noise on the deconvolved response. Further, the entire process becomes more difficult and unreliable with only a few ‘good’ pressure-buildup tests. Linearizing the convolution equation using normalized pressure or normalized rate are common alternatives for analyzing variable pressure-rate-time information. Nonetheless, Kuchuk et al. (2016) demonstrated the equivalence between the normalized pressure or normalized rate solution and Gladfelter deconvolution (Kuchuk et al. 2016) illustrated in Fig. 8, and he emphasized that this type of deconvolution is limited in applications if the flow rate varies linearly with time for simple flow geometries, and is not valid for arbitrary variable flow rates. The difficulties and uncertainties of using current deconvolution techniques undermined the critical value of integrated process using Permanent Downhole Gauges (PDG) to survey real-time data with additional data imperfections and distortions. Therefore, we propose to develop new analytic approaches as reliable diagnostic tools to interpret pressure-rate-time data using traditional PTA and RTA for unconventional reservoirs. Advantages of this analytic approach are that the algorithm: (1) employs physics-based training features and training algorithm to facilitate the PDA with the help of diagnostic plots and specialized plots; (2) enables engineers to further forecast future production rapidly and accurately with limited data; and (3) can deal with moderate noise. Additionally, the procedure is informed by physics-based training features and increasing bytes of data storage gathered from downhole. More specifically, aberrant segments, outliers, and systematic noise can be regarded as the inherent characteristics of downhole data and would likely cause interpretation bias and lead to a less representative estimation of reservoir properties. Hence, we propose a systematic workflow employing a deep-learning approach for rate-pressure reconstruction to facilitate production surveillance so that the procedure (1) inherently tolerates the effects of outliers, aberrant segments, and noise, and preserves the intrinsic characteristics during the pressure-rate-reconstruction procedure; (2) successfully generates missing production histories to fill gaps and enables entire production histories to be applicable in production analysis; and (3) robustly adapts to higher frequency data with increasing bytes cost-effectively.

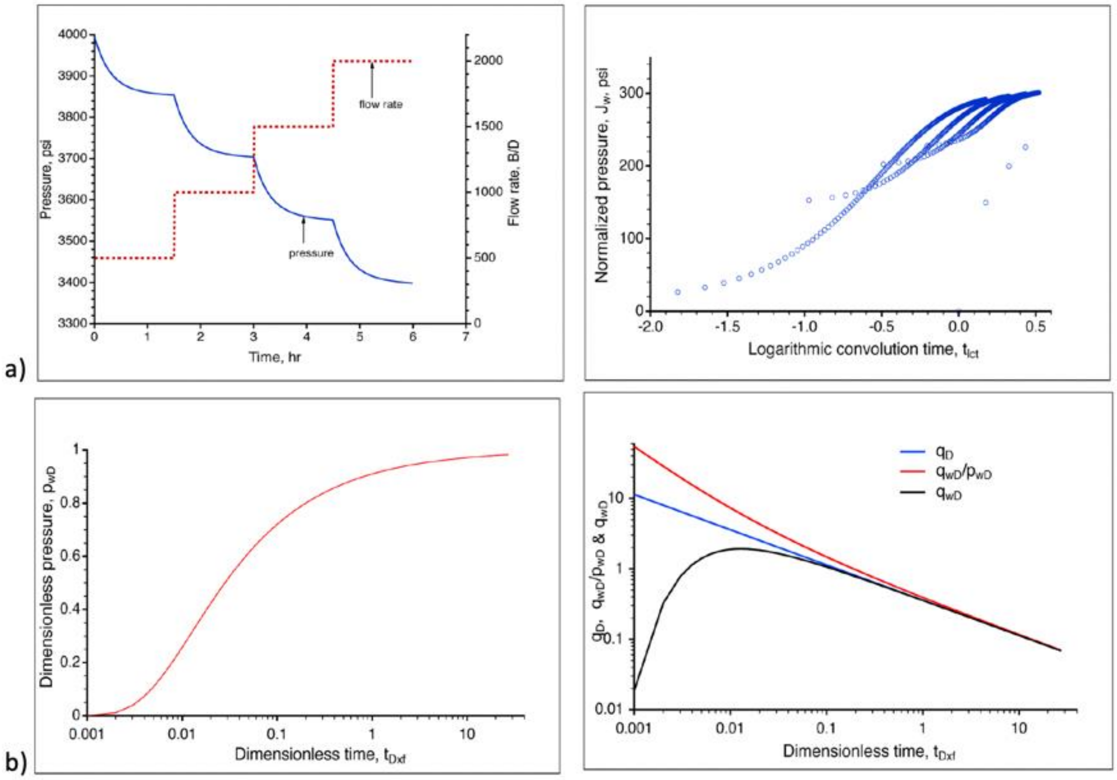


Fig. 8 (a)-Stepwise-varying flowrates, corresponding wellbore pressures, and Gladfelter deconvolution yielding a false constant-rate solution; (b)-varying wellbore pressure and normalized rate and Gladfelter deconvolution yielding a false constant-pressure solution (Kuchuk et al. 2016).

1.3 Literature Review

Deconvolution was first used to compute the equivalent constant-rate behavior of a variable rate/variable pressure system (Lee et al. 2003). von Schroeter et al. presented a time-domain deconvolution method for a specific level of noise existed in both pressures and rates by using a separable non-linear total least squares algorithm (von Schroeter et al. 2001). In addition to the significant breakthroughs presented by von Schroeter et al. (2004), Levitan derived a workflow by matching pressures and pressure derivatives from various pressure build-up periods (Levitan et al. 2003). This extended work from Levitan proved to be more practical when applied to several 'good' pressure build-up tests. Onur extended Levitan and von Schroeter's work by changing the objective function from matching the pressure drawdown to matching the pressure derivatives (Onur et al. 2012). Benefitting from Onur's method, the dependency of initial pressure can be removed and the initial pressure estimation error impact on late-time periods deconvolved response can be alleviated. This methodology is based on pressure transient analysis for conventional reservoirs in which constant rate drawdown-pressure responses are acquired to determine flow regimes and estimate reservoir properties. In unconventional reservoirs, rate transient analysis is commonly used as a substitute for pressure transient analysis. Constant-pressure decline curves are used to diagnose decline types. Kuchuk et al. proposed rate/pressure deconvolution to calculate the constant pressure drawdown response (Kuchuk et al. 2005). As previously stated, piecewise pressure control or piecewise rate control as the wellbore inner condition in this research cannot be transformed to RNP or

PNR since flow rates vary nonlinearly with time. Thus, our analytic methods have unique value as real-world data are usually correlated nonlinearly.

The issues in the methods discussed above used in PTA were addressed by Liu and Horne (2013). For example, they advised that only part of the ‘good’ data should be used to ensure a consistent deconvolution process. Noise deteriorates the deconvolution results especially for rate/pressure deconvolution methodology (Kuchuk et al. 2005). Moreover, the assumption of reasonable accuracy in identifying breakpoints in flow regimes is sometimes unrealistic. To make full use of the entire history data, Liu and Horne proposed a data-mining approach using kernelization (Liu and Horne 2013). The key idea in their ‘kernel tricks’ is to use the superposition principle; the reservoir pressure response is controlled by a linear combination of individual flow rate described in Fig. 9. They demonstrated that their non-parametric data-mining algorithm does not require any prior knowledge about the physical geological model, and that the appropriate reservoir model can be approximated in high-dimensional polynomial space with corresponding physics-based training features. This modern concept of combining data-driven and model-based approaches is critical and should become repeatable workflow in many engineering analyses (Deng et al. 2018).

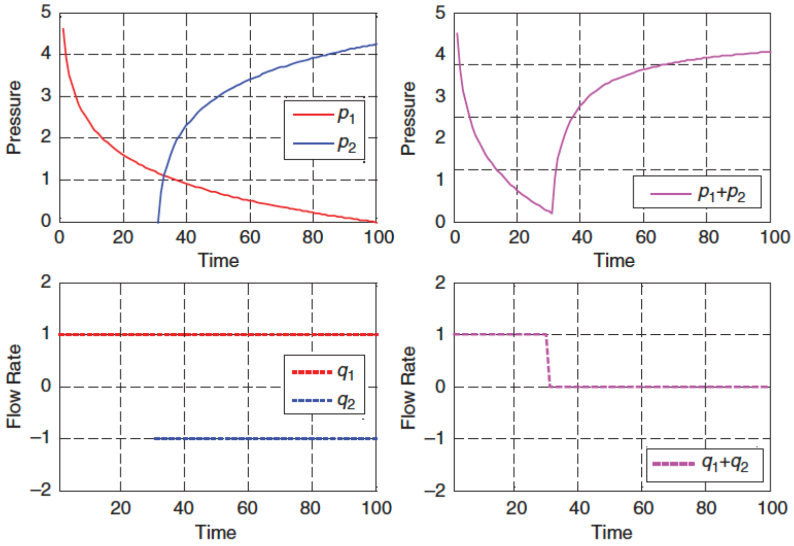
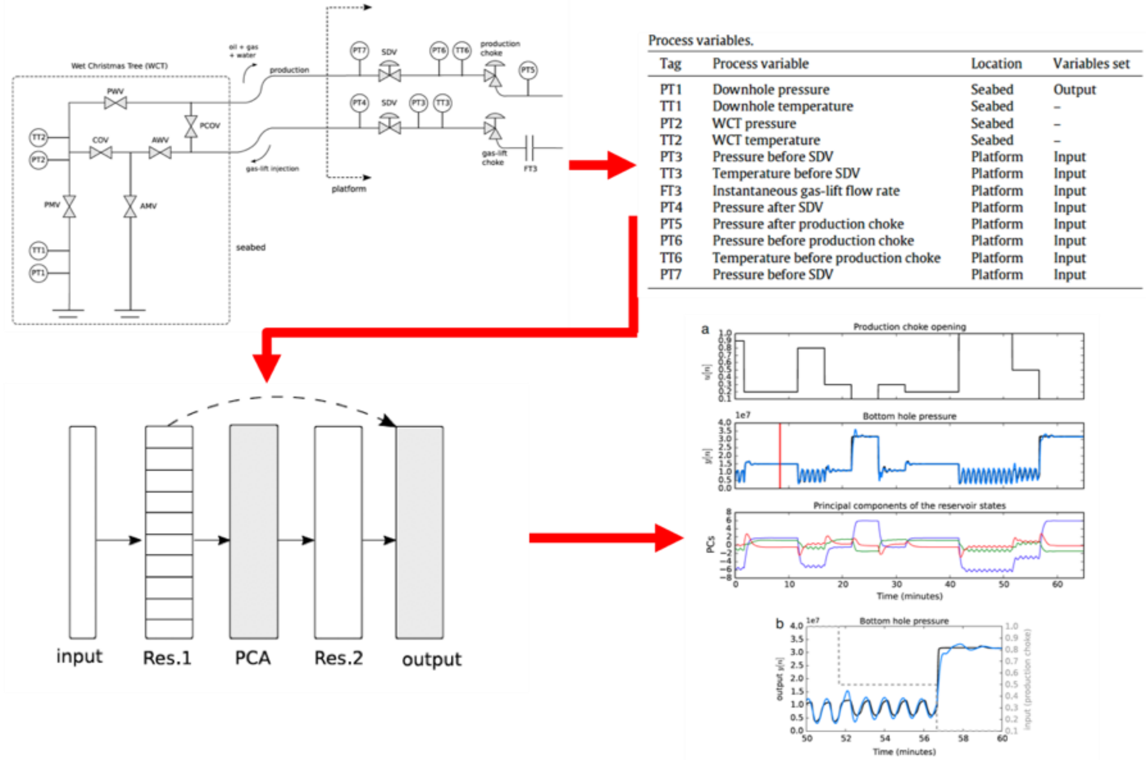


Fig. 9-Demonstration of superposition principle (Liu et al. 2013).

In this study, we first extended Liu and Horne’s work; however, instead of solving pressure-rate deconvolution using ‘kernel tricks’, we adopted an echo state network (ESN) approach to solve both pressure/rate and rate/pressure deconvolution problems. The echo state network is a novel design for non-parametric modeling of nonlinear dynamic systems. Unlike traditional recurrent neural networks (RNNs), it contains a large sparsely connected reservoir where only the connections from the reservoir to the output layer are trainable by linear regression algorithms, and other connected weights such as reservoir internal weights and connections between inputs and reservoirs remain unchanged. A recent application of ESN in the oil and gas industry was to estimate downhole pressure in gas-lift oil wells (Antonelo et al. 2017). The workflow is shown in Fig. 10. Eight gauge recordings deployed on the surface with the choke opening versus time were used as inputs to estimate downhole pressure. Prospectively, the critical concept is trying to set up a proxy model to substitute some empirical mathematical correlation such as the Beggs and

Brill correlation. Yet the connection between physics and the data-driven solution still remains veiled. Pressure variation with time is a chaotic time-series problem, and the ESN approach proved to offer an effective approximation of model-free prediction of large spatiotemporal chaotic systems (Pathak et al. 2018). Additionally, Hermans et al. demonstrated that an infinite number of hidden states which employ large, randomly initiated neural networks (for instance, ESN) can be considered equivalent to recursive kernels (Hermans and Schrauwen, 2012). However, Han and Xu indicated that the hidden states could be ill-posed and their singular values could possibly decay to zero when the number of real-time training samples is less than the number of hidden states (Han and Xu 2018). To overcome this problem, the Laplacian Eigenmaps Coupled Echo-State Network (LEESN) was utilized to create a lower dimensional representation of relatively large hidden states, and this algorithm intuitively solves the production forecasting problem in unconventional reservoirs during the early production stage when fewer months of production data are available. Additionally, the Laplacian Eigenmap is relatively insensitive to outliers and noise and is more robust with real-time data.



Process variables.

Tag	Process variable	Location	Variables set
PT1	Downhole pressure	Seabed	Output
TT1	Downhole temperature	Seabed	–
PT2	WCT pressure	Seabed	–
TT2	WCT temperature	Seabed	–
PT3	Pressure before SDV	Platform	Input
TT3	Temperature before SDV	Platform	Input
FT3	Instantaneous gas-lift flow rate	Platform	Input
PT4	Pressure after SDV	Platform	Input
PT5	Pressure after production choke	Platform	Input
PT6	Pressure before production choke	Platform	Input
TT6	Temperature before production choke	Platform	Input
PT7	Pressure before SDV	Platform	Input

Fig. 10-Echo-State Networks for downhole pressure estimation (Antonelo et al. 2017).

Additionally, Tian and Horne extended their previous work using various data-mining tricks and performed sensitivity analysis (Tian and Horne, 2015). The applicability of using physics-based training features for production surveillance was addressed in their paper; however, few discussions have been presented explaining how engineers can implement this machine learning algorithm in the real world. These considerations propelled us to investigate an in-house workflow using a deep-learning approach for real-time data integrity. In the past few years, permanent downhole gauges have been equipped for real-time monitoring. Common production data management strategies are established in a rigorous procedure to ensure a Quality Assurance (QA)/Quality Check (QC) of structured and unstructured data (Gonzalez et al. 2018). Unexpected failure, system offline and random sampling error can cause false real-time data such as missing values, aberrant segments and outliers. Classical models rely on the ‘linear interpolation’ scheme such as autoregressive-moving average model (ARMA) and autoregressive-integrated moving average model (ARIMA) aiming to smoothly fit observations (Ansley et al. 1984). However, ignorance of the correlation of variables over time could be critical and possibly result in an erratic interpolation (Che et al. 2018). Aberrant segments, outliers, and systematic noise can be regarded as inherent behavior of downhole data and would highly likely trigger interpretation bias and lead to a less representative estimation of reservoir properties. Recently, Madasu utilized LSTM to correlate actual pumping slurry rates and surface treatment pressures to monitor the hydraulic fracturing process depicted in Fig. 11 (Madasu, 2018). Another intriguing study illustrated in Fig. 12 was conducted by Zhang, who demonstrated the Cascaded Long Short-Term Memory (CLSTM) method (the design of cascaded combination of LSTM units and dense units) as an accurate and cost-effective way to reconstruct well-logging curves by comparing the performances of CLSTM and Fully Connected Neural Networks (FCNNs) (Zhang, et al. 2018). Still, no detail explanation of the physics behind the algorithm was provided and the outstanding performance still remains a mystery. Thus, in the second part of our study, we employed the basic idea from Zhang by using CLSTM for filling the missingness and adapted the algorithm for incomplete production history reconstruction using physics-based training features.

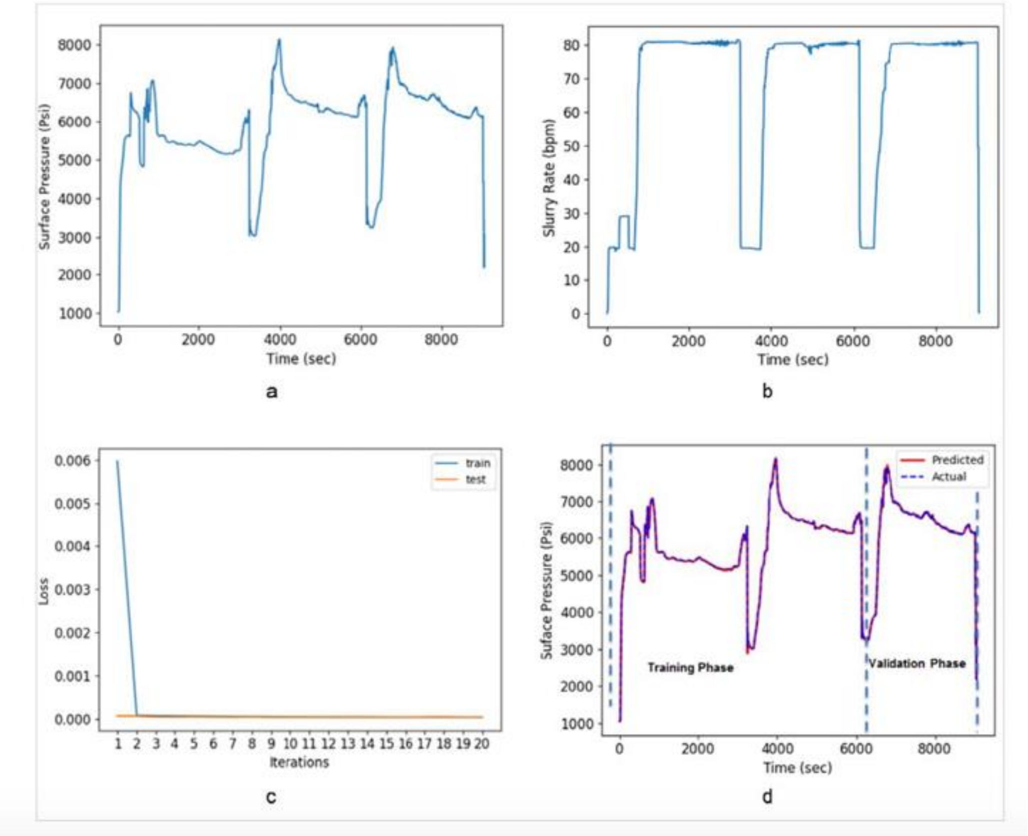


Fig. 11-One example using LSTM and available slurry rate to predict surface pressure and monitor hydraulic fracturing process (Madasu et al. 2018).

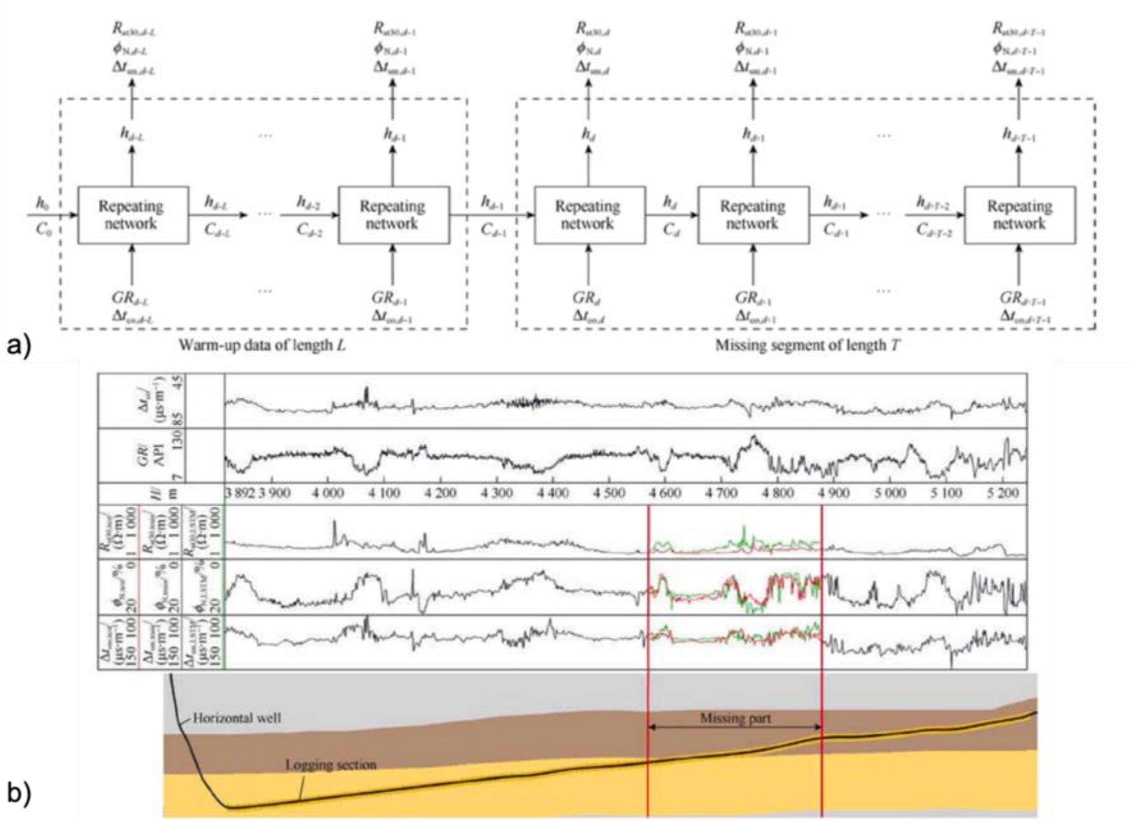


Fig. 12 (a)-CLSTM schematic topology for missing well-logging curve reconstruction; (b)-A field example using CLSTM to reconstruct missing logging curves (Zhang et al. 2018).

CHAPTER II

METHODOLOGY

2.1 Echo State Network

Our pilot works successfully demonstrated that combining the ‘kernel trick’ method and physics-based training features broadens the ‘pathway’ to deconvolve variable pressures and variable rates into a constant-rate pressure-drawdown response using relatively more complete pressure-rate-time information. However, choosing the correct kernel function could be arbitrary and is usually achieved by trial and error. Additionally, the kernel space depends on the feature space, and the feature space especially using superposition in time, could cause a large-size matrix that leads to computationally expensive process. Echo State Network constrains the reservoir space as an approximation of kernel space to improve the computation expenses and instead of randomly selecting a kernel function, the reservoir space is controlled by an extra hyperparameter.

Part of section 2.1 is reprinted with permission of the Society of Petroleum Engineers from Pan, Y., Zhou, P., Deng, L., Lee, J. 2019. Production Analysis and Forecasting for Unconventional Reservoirs Using Laplacian Echo-State Networks. Proceedings of SPE Western Regional Meeting, San Jose, California, 23-26 April. SPE-195243-MS. DOI:10.2118/195243-MS, and from Pan, Y., Deng, L., Lee, J. 2019. Data-Driven Deconvolution Using Echo-State Networks Enhances Production Data Analysis in Unconventional Reservoirs. Proceedings of SPE Eastern Regional Meeting, Charleston, West Virginia, 15-17 October. SPE-196598-MS. DOI: 10.2118/196598-MS. Permission conveyed through Copyright Clearance Center, Inc.

2.1.1 Background of Echo State Network

Considerable research on neural networks, especially on recurrent neural networks (RNNs), has proved the capability of supervised-learning from data to establish feature-based classifiers and nonlinear predictive models. However, the nonconvex objective functions which result in the costly computation of gradients make it infeasible for applications (Scardapane and Wang 2017). An alternative approach known as reservoir computing (RC) under the paradigm of RNN was then developed. The key characteristics are first to use randomization to define a feature-based, commonly data-dependent mapping, and then by transforming the feature inputs into a higher dimensional 'reservoir' space which is comparable with Hilbert space or Kernel space, and the learning process is ultimately simplified (Dambre et al. 2012). Second, the optimization of the objective function becomes a standard linear-square adaptation problem. This idea initiated the twin pair of echo state networks (ESNs) (Jaeger and Haas 2004) and liquid state machines (LSMs) (Maass and Markram 2002). In this research, we focus on the ESN, which is made of three layers, the input layer, the reservoir layer which is utilized with random initialization and finally the readout layer which can be linearly adaptable (Scardapane and Wang 2017). From the topology as shown in Fig. 13, after the transformation of feature inputs through the weighted connection between input and 'reservoir,' each neuron in the reservoir acts as a dynamic filter/extractor to create the chronological 'echo' state of the input sequence for classification/regression purposes. The basic insight behind the application of ESNs is the intriguing notions connecting the reservoir (infinite neurons) and recursive kernel methods that have been proved by Hermans and Schrauwen (2012).

For those reasons, it is crucial to creating physics-based training features to successfully train the algorithms and perform further deconvolution in production data analysis and production forecasting. A handful of global parameters could also influence the model performance, thus these must be treated carefully in terms of model stability and locality (Lukoševičius, 2012). However, this was beyond the scope of this study. Lukoševičius's paper was adopted as a practical guide of applying ESN. During experiments, we found once these global values are set to fulfill the stability check, the model performance depends solely on the dimensionality of the reservoir space with the proposed physics-based training features.

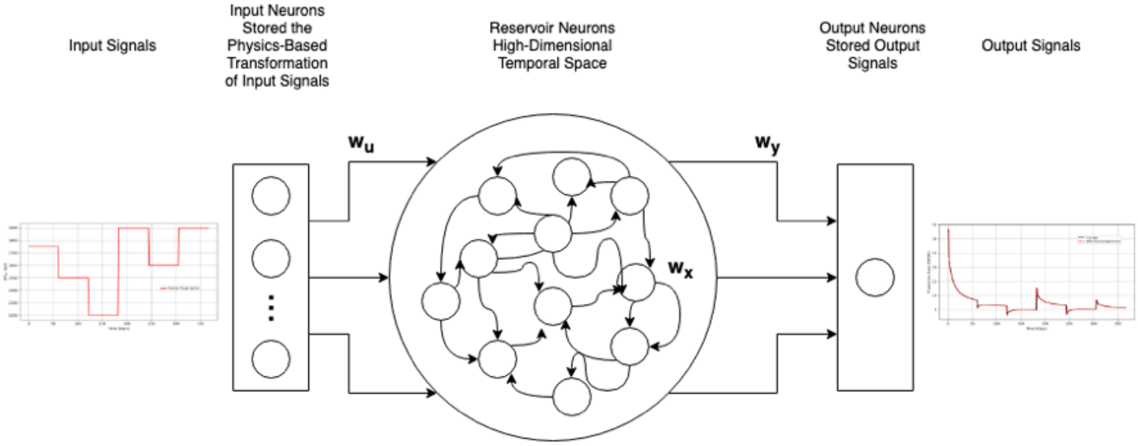


Fig. 13-The topology of a typical Echo State Network (Modified from Jaeger and Haas 2004).

The specific category of machine learning algorithm presented here is supervised learning for regression, in which for a given dataset containing measured target values y and feature values u, the ultimate purpose is to correlate $\mathbf{y}$ and $\mathbf{u}$ to approximate a general function $\mathbf{y} = f(\mathbf{u})$ that could be utilized for predictions. Although this type of algorithm has been widely applied in engineering problems, there is still a special case in which the targets $\mathbf{y}(t)$ and features $\mathbf{u}(t)$ are both functions of time that has not been explored thoroughly. A more rigorous approach is to introduce an additional variable t and then incorporate it as a sequential pattern into the machine learning algorithm. Thereafter, the reservoir computing under the paradigm of Recurrent Neural Networks (RNNs) emerged by considering the inherent chronological behavior inside the data sequence. In this chapter, we mainly focus on Echo State Networks (ESNs) - one of kind in reservoir computing. It was demonstrated as an approximation of recursive kernels Support Vector Machine (SVM) and can be expressed as following (Hermans and Schrauwen 2012):

$$\mathbf{y}(t) = \sum_{i=1}^{\mathrm{N}} \alpha_i \kappa_t\big(\mathbf{u}(t\colon t-\tau), \mathbf{u_i}(0\colon -\tau)\big) + \beta \tag{1}$$

$$\mathbf{y}(t) = \sum_{i=1}^{\mathrm{N}} \alpha_i \mathbf{x_i} \cdot \mathbf{x}(t) + \beta = \boldsymbol{\theta}^{\mathbf{T}} \cdot \mathbf{x}(t) + \beta \tag{2}$$

where $\mathbf{x_i}$ is the last hidden state caused by time series $\mathbf{u}(0\colon -\tau)$, and $\mathbf{x}(t)$ is the hidden state triggered by $\mathbf{x}(t\colon -\infty)$. A similar idea was borrowed from 'kernel tricks'; the kernel function can be expressed as $\kappa\big(\mathbf{x_i}, \mathbf{x}(t)\big) = \mathbf{x_i} \cdot \mathbf{x}(t)$. Superior to the recursive kernel tricks, ESNs introduces a parametric control of kernel space.

2.1.2 Leaky Integrator Echo State Network

In the recent development of ESNs, the Leaky Integrator Echo State Network (LIESN) was first developed and demonstrated mathematically by Jaeger (Jaeger et al. 2007). Thereafter, Lukoševičius came up with a practical guide to applying LIESN for various engineering problems (Lukoševičius 2012). The basic LIESN method can be summarized with the following mathematical equations:

$$\tilde{\mathbf{x}}(t) = \varphi\big(\mathbf{W_u}[1, \mathbf{u}(t)] + \mathbf{W_x}\mathbf{x}(t-1)\big) \tag{3}$$

$$\mathbf{x}(t) = (1-\alpha)\mathbf{x}(t-1) + \alpha\tilde{\mathbf{x}}(t) \tag{4}$$

$$\mathbf{y}(t) = \mathbf{W_y}[1; \mathbf{u}(t); \mathbf{x}(t)] \tag{5}$$

where φ is the activation function, which can be sigmoid, logistic, or ReLu etc., however, in our study, we adopted the hyperbolic tangent function:

$$\varphi(x) = \tanh(x) = \left(\frac{1-e^{-2x}}{1+e^{-2x}}\right) \tag{6}$$

The objective function is typically a Mean-Square Error (MSE) function:

$$E(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{\mathrm{N_y}}\sum_{i=1}^{\mathrm{N_y}}\left(\frac{1}{\mathrm{N_t}}\sum_{t=1}^{\mathrm{N_t}}\big(\mathbf{y_i}(t) - \hat{\mathbf{y}}_\mathbf{i}(t)\big)^2 + \beta_r\|\mathrm{w}_{y_i}\|^2\right) \tag{7}$$

In the training procedure, the state matrix can be described as below:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x^T(1)} \\ \mathbf{x^T(2)} \\ \vdots \\ \mathbf{x^T(N_t)} \end{bmatrix} = \begin{bmatrix} \mathrm{x}_1(1) & \mathrm{x}_2(1) & \cdots & \mathrm{x}_{\mathrm{N_r}}(1) \\ \mathrm{x}_1(2) & \mathrm{x}_2(2) & \cdots & \mathrm{x}_{\mathrm{N_r}}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{x}_1(\mathrm{N_t}) & \mathrm{x}_2(\mathrm{N_t}) & \cdots & \mathrm{x}_{\mathrm{N_r}}(\mathrm{N_t}) \end{bmatrix} \tag{8}$$

Similarly, the target matrix can be written in a matrix form shown below:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}(1) \\ \mathbf{y}(2) \\ \vdots \\ \mathbf{y}(\mathrm{N_t}) \end{bmatrix} \tag{9}$$

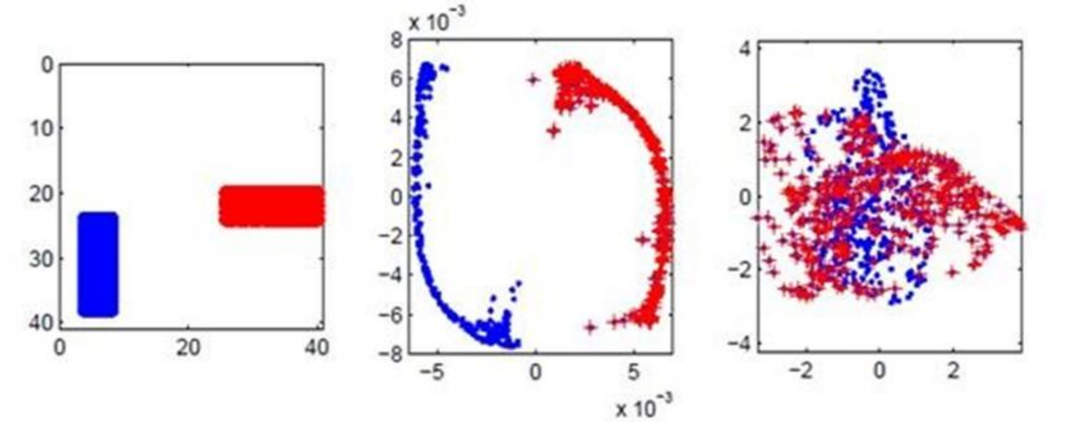
The universal stable solution to Eq. 7 is ridge regression also known as regression with Tikhonov regularization:

$$\mathbf{W_y} = \mathbf{Y}\mathbf{X^T}(\mathbf{X}\mathbf{X^T} + \beta_r\mathbf{I})^{-1} \tag{10}$$

Ridge regression has been intensively applied to cope with the overfitting issue. The tuning parameter $\beta_r > 0$ has been widely studied and we did not devote much attention to it in this study. Usually, a K-fold (e.g., K=5 or 10) cross-validation is used to select the appropriate β_r and this procedure has been well explained fully in various studies (Tian and Horne 2015). An alternative is to use a Gaussian process/Bayesian Regression interpretation of the linear readout for setting β_r directly which will be introduced in section 2.1.4 (Lukoševičius 2012).

2.1.3 Laplacian Eigenmaps Coupled Echo State Network

Laplacian Eigenmaps (LE) was proposed by Belkin, et al. to provide a proper representation of complex data by transforming data in high-dimensional space into a low-dimensional manifold (Belkin and Niyogi 2003). Unlike Principle Component Analysis (PCA), this method explicitly considers the structure of the manifold on which data could possibly locate. A benchmark result comparing LE and PCA is illustrated in Fig. 14.



Original Data LE Transformation PCA Transformation

Fig. 14-Benchmark dimension reduction problem using Laplacian Eigenmaps and Principal Component Analysis (Belkin and Niyogi 2003).

LE was first coupled with ESN by Han and Xu, and it was able to extract the nonlinear manifold from the chaotic time-series data (Han and Xu 2018). Additionally, the LE is relatively insensitive to outliers and noise. The coupled algorithm is mathematically interpreted below:

Please recall the standard leaky integrator ESNs' dynamic system from Eq. 3 to Eq. 5, modified slightly for simplicity:

$$\tilde{\mathbf{x}}(t) = \varphi\left(\mathbf{W}_{\mathbf{u}}\mathbf{u}(t) + \mathbf{W}_{\tilde{\mathbf{x}}}\tilde{\mathbf{x}}(t-1)\right) \tag{11}$$

$$\mathbf{y}(t) = \mathbf{W}_{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{x}}(t) \tag{12}$$

The schematic topology of LEESN is presented in Fig. 15. To enhance the consistency of implementation of the proposed LEESN algorithm, we present a workflow for analysis of data with moderate noise and highly variable production histories; this workflow enables engineers to recognize flow patterns and estimate reservoir properties.

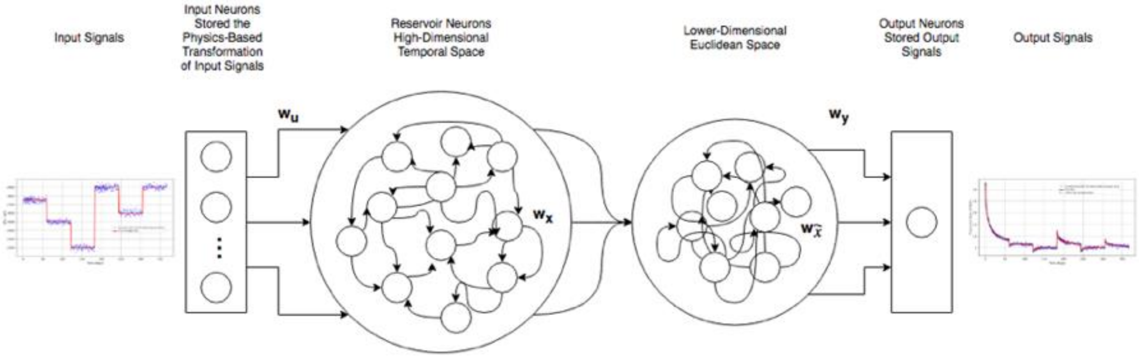


Fig. 15-The topology of a Laplacian Eigenmaps Coupled Echo State Network.

Deconvolution Process

1. Prepare the training features.
2. Standardize or normalize the inputs and outputs.
3. Initialize ESN reservoir parameters; e.g., neuron numbers in reservoir $\mathrm{N_r}$.
4. Activate reservoir neurons and acquire reservoir states $\mathbf{X}$ using training samples flushing through Eqs. 3 and 4.
5. For $\mathcal{H} = 1\colon \mathrm{N_r}$ do

   Construct an adjacency matrix based on the reservoir states $\mathbf{X}$ using K-Nearest-Neighbor (KNN) and then sort them from smallest to largest in terms of the distance value d_{st}.

   The reservoir state matrix is expressed as $\mathbf{X} = [\mathbf{X_1}\ \mathbf{X_2} \ldots \mathbf{X_{N_r}}]$, where $\mathbf{X_i}$ indicates $[\mathbf{x_i}(1)\ \mathbf{x_i}(2) \ldots \mathbf{x_i}(\mathrm{N_t})]$.

$$d_{st} = \sqrt{\sum_{j=1}^{\mathrm{N_r}} \left(\mathbf{X_i} - \mathbf{X_j}\right)^2} \tag{13}$$

   d_{st} is then sorted from smallest to largest: $z^1 < z^2 < \cdots < z^{\mathrm{N_r}}$.

Choose the appropriate value to connect the edge of the adjacency matrix by using the heat kernel:

$$W_{ij} = e^{-\|\mathbf{x_i} - \mathbf{x_j}\|^2/\sigma} \tag{14}$$

where σ is the predefined kernel width and $W_{ij} = 0$ when the two elements are not connected in the adjacency matrix.

Calculate the eigenmaps and minimize the objective function:

$$argmin \sum_{i,j} \left(\mathbf{f_i} - \mathbf{f_j}\right)^2 W_{ij} \equiv argmin \frac{1}{2} \sum_{i,j} \left(\mathbf{f_i} - \mathbf{f_j}\right)^2 W_{ij} = argmin\ \mathbf{f^T} L \mathbf{f} \tag{15}$$

A generalized eigenvalue problem is eventually formed. Determine the minimum eigenvalue solution.

$$L\mathbf{f} = \lambda \mathbf{D}\mathbf{f} \tag{16}$$

Denote $\left[\mathbf{f_0}, \mathbf{f_1}, \mathbf{f_2}, \dots, \mathbf{f_{N_r - 1}}\right]$ as the ordered solutions of Eq. 16 according to the eigenvalues:

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{\mathrm{N_r}-1} \tag{17}$$

$$\boldsymbol{L}\mathbf{f_0} = \lambda_0 \mathbf{D}\mathbf{f_0}, \boldsymbol{L}\mathbf{f_1} = \lambda_1 \mathbf{D}\mathbf{f_1}, \dots, \quad \boldsymbol{L}\mathbf{f_{N_r - 1}} = \lambda_{\mathrm{N_r}-1} \mathbf{D}\mathbf{f_{N_r - 1}} \tag{18}$$

The first eigenvector $\mathbf{f_0}$ is neglected corresponding to the eigenvalue $\lambda_0 = 0$. The next $\mathcal{H}$ eigenvectors are the lower transformation factors $\mathbf{F} = [\mathbf{f_1}, \mathbf{f_2}, \dots, \mathbf{f}_{\mathcal{H}}]$.

$$\dot{\mathbf{X}} = \mathbf{F^T}\widetilde{\mathbf{X}} \tag{19}$$

$$\dot{\mathbf{W}}_\mathbf{y} = \mathbf{Y}\dot{\mathbf{X}}^\mathbf{T}\left(\dot{\mathbf{X}}\dot{\mathbf{X}}^\mathbf{T} + \boldsymbol{\beta}_r \mathbf{I}\right)^{\mathbf{-1}} \tag{20}$$

$$\widehat{\mathbf{Y}} = \dot{\mathbf{W}}_\mathbf{y}^\mathbf{T}\dot{\mathbf{X}} \tag{21}$$

Using the Hannan-Quinn Criterion to select best $\mathcal{H}$,

$$\boldsymbol{\epsilon} = \widehat{\mathbf{Y}} - \mathbf{Y} \tag{22}$$

$$\mathrm{HQ}(\mathcal{H}) = -\ln \widehat{\sigma}^2 + \frac{2\mathcal{H}\ln\left(\ln\left(\mathrm{N_y}\right)\right)}{\mathrm{N_y}} \tag{23}$$

where $\widehat{\sigma}^2 = \frac{1}{\mathrm{N_y}}\sum_{n=1}^{\mathrm{N_y}} \lVert\epsilon\rVert^2$.

6. Given constant $q(t)$, repeat steps 1 to 4 to generate $\mathbf{u}^{\star}$.
7. Using the saved $\mathcal{H}$ and $\dot{\mathbf{W}}_{\mathbf{y}}^{\mathbf{T}}$,

$$\dot{\mathbf{X}}^{*} = \mathbf{F}^{\mathbf{T}}\widetilde{\mathbf{X}}^{*} \tag{24}$$

$$\widehat{\mathbf{Y}}^{*} = \dot{\mathbf{W}}_{\mathbf{y}}^{\mathbf{T}}\dot{\mathbf{X}}^{*} \tag{25}$$

$\widehat{\mathbf{Y}}^{*}$ is the deconvolution response.

Using the lower dimension of $\dot{\mathbf{X}}^{\mathbf{T}}\dot{\mathbf{X}}$ could solve the ill-conditioned problem as stated previously. It also reduces the computational expenses to $\mathrm{o}(\mathcal{H}^3)$ from $\mathrm{o}(\mathrm{N_r^3})$. Again, lower transformation $\dot{\mathbf{X}}$ tolerates noise and outliers due to its locality preserving attributes. The applications are performed in the section 3.1 and section 3.2.

2.1.4 Bayesian Echo State Network

The original solver for Leaky Integrator Echo State Network (LIESN) and Laplacian Echo State Network (LEESN) is the linear solver – ridge regression expressed in Eq. 10. An alternative Bayesian regression can substitute for ridge regression to automatically estimate model parameters using only training data and additionally provide the confidence interval of prediction. However, Bayesian regression assumes that errors are zero-mean Gaussian noise with the variance β_{BR} and the likelihood function can be expressed as following (Li et al. 2012):

$$p(\mathbf{y}|\mathbf{W_y},\beta_{BR})=\left(\frac{\beta_{BR}}{2\pi}\right)^{\frac{\mathrm{N}}{\mathbf{2}}}\exp\left\{-\frac{\beta_{BR}}{2}\left(\mathbf{y}-\mathbf{XW_y}\right)^2\right\} \tag{26}$$

The prior probability over the output weights $\mathbf{W_y}$ is governed by α_{BR}:

$$p(\mathbf{W_y}|\alpha_{BR})=\left(\frac{\alpha_{BR}}{2\pi}\right)^{\frac{\mathrm{N_r}}{\mathbf{2}}}\exp\left\{-\frac{\alpha_{BR}}{2}\mathbf{W_y^T W_y}\right\} \tag{27}$$

Combining Eq. 26 and Eq. 27, the posterior distribution over output weights can be acquired according to Bayes' rule. The Gaussian prior $p(\mathbf{W_y}|\alpha_{BR})$ is a Gaussian distribution with mean and variance as below:

$$\kappa=\beta_{BR}\Sigma\mathbf{X^T}\boldsymbol{y} \tag{28}$$

$$\Sigma=(\alpha_{BR}\mathbf{I}+\beta_{BR}\mathbf{X^T X})^{-1} \tag{29}$$

Model performance depends on the critical hyper parameter α_{BR} and β_{BR}, which can be obtained from

$$\alpha_{BR}=\frac{\gamma}{\kappa^T\kappa} \tag{30}$$

$$\beta_{BR}=\frac{(\mathrm{N_t}-\gamma)}{\left\|\mathbf{y}-\widetilde{\mathbf{X}}\kappa\right\|^2} \tag{31}$$

where γ is defined as:

$$\gamma=\sum_{i=1}^{\mathrm{N_r}}\frac{\lambda_{BR}}{\lambda_{BR}+\alpha_{BR}} \tag{32}$$

λ_{BR} is the eigenvalue of the Hessian matrix of the error function shown as below (Li et al. 2015):

$$E = \frac{1}{2}\left(\mathbf{y} - \widetilde{\mathbf{X}}\mathbf{W}_{\mathbf{y}}\right)^2 \tag{33}$$

The ultimate goal is to minimize the equation

$$F\left(\mathbf{W}_{\mathbf{y}}\right) = \frac{\beta_{BR}}{2}\left(\mathbf{y} - \mathbf{X}\mathbf{W}_{\mathbf{y}}\right)^2 + \frac{\alpha_{BR}}{2}\left(\mathbf{W}_{\mathbf{y}}^{\mathbf{T}}\mathbf{W}_{\mathbf{y}}\right) \tag{34}$$

2.2 Long Short-Term Memory

With more data gathering from downhole leading to increasing bytes in data storage, the computation expenses in kernel space or equivalent reservoir space can be increased drastically. Therefore, we adopted a deep learning approach, Long-Short Term Memory (LSTM) considering the sequential dependencies to reconstruct the missing production rates with the presence of aberrant segments, outliers and systematic noise.

Part of section 2.2 is reprinted with permission of URTeC from Pan, Y., Bi, R., Zhou, P., Deng L., Lee, J. 2019. An Effective Physics-Based Deep Learning Model for Enhancing Production Surveillance and Analysis in Unconventional Reservoirs. Proceedings of Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July. URTEC-2019-145. DOI:10.15530/urtec-2019-145. Copyright 2019 URTeC.

2.2.1 Long Short-Term Memory Unit

A typical LSTM unit process comprises two components, the forward components and the backward components. The topology of LSTM unit is depicted in Fig. 16. The critical improvements beyond the SRNs are the introduction of the forget gate (Gers et al. 1999) and the utilization of full gradient backpropagation through time (BPTT) (Hochreiter et al. 1997). The forget gate allows learning for continual tasks such as language processing. BPTT makes gradient check feasible in practical implementation and alleviates gradient vanishing, gradient exploding problems (Greff et al. 2017).

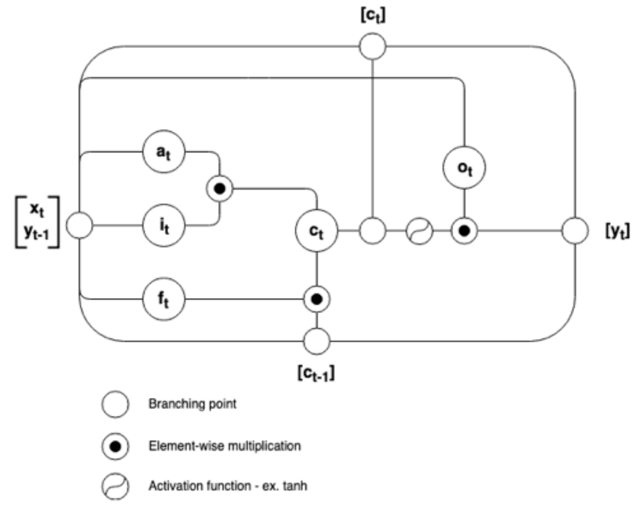


Fig. 16-The topology of a typical LSTM unit without peephole connection (Modified from Olah 2015).

Forward components: Mathematically, the gates are defined as follows: input gate i_t, forget gate f_t, output gate o_t, the formula for each gate is presented below.

Block input activation:

$$\bar{a}_t = W_a \cdot x_t + U_a \cdot y_{t-1} + b_a \tag{35}$$

$$a_t = h(\bar{a}_t) \tag{36}$$

Input gate:

$$\bar{i}_t = W_i \cdot x_t + U_i \cdot y_{t-1} + b_i \tag{37}$$

$$i_t = \sigma(\bar{i}_t) \tag{38}$$

Forget gate:

$$\bar{f}_t = W_f \cdot x_t + U_f \cdot y_{t-1} + b_f \tag{39}$$

$$f_t = \sigma(\bar{f}_t) \tag{40}$$

Output gate:

$$\bar{o}_t = W_o \cdot x_t + U_o \cdot y_{t-1} + b_o \tag{41}$$

$$o_t = \sigma(\bar{o}_t) \tag{42}$$

Updating iteratively and passing internal states:

$$c_t = a_t \odot i_t + f_t \odot c_{t-1} \tag{43}$$

$$y_t = h(c_t) \odot o_t \tag{44}$$

x_t is the input vector at time t, N is the number of LSTM units, and M is the number of inputs.

- Input weights: $W_a, W_i, W_f, W_o \in \mathfrak{R}^{N \times M}$.
- Recurrent weights: $U_a, U_i, U_f, U_o \in \mathfrak{R}^{N}$.
- Bias weights: $b_a, b_i, b_f, b_o \in \mathfrak{R}^{N}$.

σ, h are the nonlinear activation functions -- more specifically in this study, σ is the sigmoid function $\left(\sigma(x) = \frac{1}{1+e^{-x}}\right)$, h is the tanh function $\left(\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}\right)$. In the above equations, $\odot$ represents the pointwise multiplication of two vectors and inner productions are denoted as $\cdot$.

Backward Components: After the forward components are fully calculated, the backpropagation through time (BPTT) should be performed to back transfer the errors and the weights and states are updated. To simplify the equation, we define

$$G_t = \begin{bmatrix} a_t \\ i_t \\ f_t \\ o_t \end{bmatrix}, W = \begin{bmatrix} W_a \\ W_i \\ W_f \\ W_o \end{bmatrix}, U = \begin{bmatrix} U_a \\ U_i \\ U_f \\ U_o \end{bmatrix}, b = \begin{bmatrix} b_a \\ b_i \\ b_f \\ b_o \end{bmatrix} \tag{45}$$

In the BPTT

$$\delta y_t = \Delta_t + \Delta y_t \tag{46}$$

$$\delta c_t = \delta y_t \odot o_t \odot h'(c_t) + \delta c_{t+1} \odot f_{t+1} \tag{47}$$

$$\delta b = \sum_{t=0}^{N_t} \delta G_{t+1} \tag{48}$$

$$\delta a_t = \delta c_t \odot i_t \odot h'(\bar{a}_t) \tag{49}$$

$$\delta i_t = \delta c_t \odot a_t \odot \sigma'(\bar{\imath}_t) \tag{50}$$

$$\delta f_t = \delta c_t \odot c_{t-1} \odot \sigma'(\bar{f}_t) \tag{51}$$

$$\delta o_t = \delta y_t \odot h(c_t) \odot h'(\bar{o}_t) \tag{52}$$

$$\delta x_t = W^T \cdot \delta G_t \tag{53}$$

$$\Delta y_t = U^T \cdot \delta G_{t+1} \tag{54}$$

Here, Δ_t is the vector of the difference passed down from the layer above. Δy_t is the output difference computed by the next time-step. Finally, the gradients for the weights are updated as follows:

$$\delta W = \sum_{t=0}^{N_t} \delta G_t \otimes x_t \tag{55}$$

$$\delta U = \sum_{t=0}^{N_t - 1} \delta G_{t+1} \otimes y_t \tag{56}$$

$$\delta b = \sum_{t=0}^{N_t} \delta G_{t+1} \tag{57}$$

2.2.2 Denoising Long Short-Term Memory

The Denoising Long Short-Term Memory (DeLSTM) approach investigated in this research is based on the structure of deep bidirectional Long Short-Term Memory (DBLSTM) for voice conversion proposed by Wu (Wu et al. 2017). Recall that, despite the purpose of denoising the target signals, one of our main tasks is to reconstruct the missing and/or distorted signals for the scenario in which only production rates are recorded. Therefore, the training scheme is modified and generalized as shown in Fig. 17.

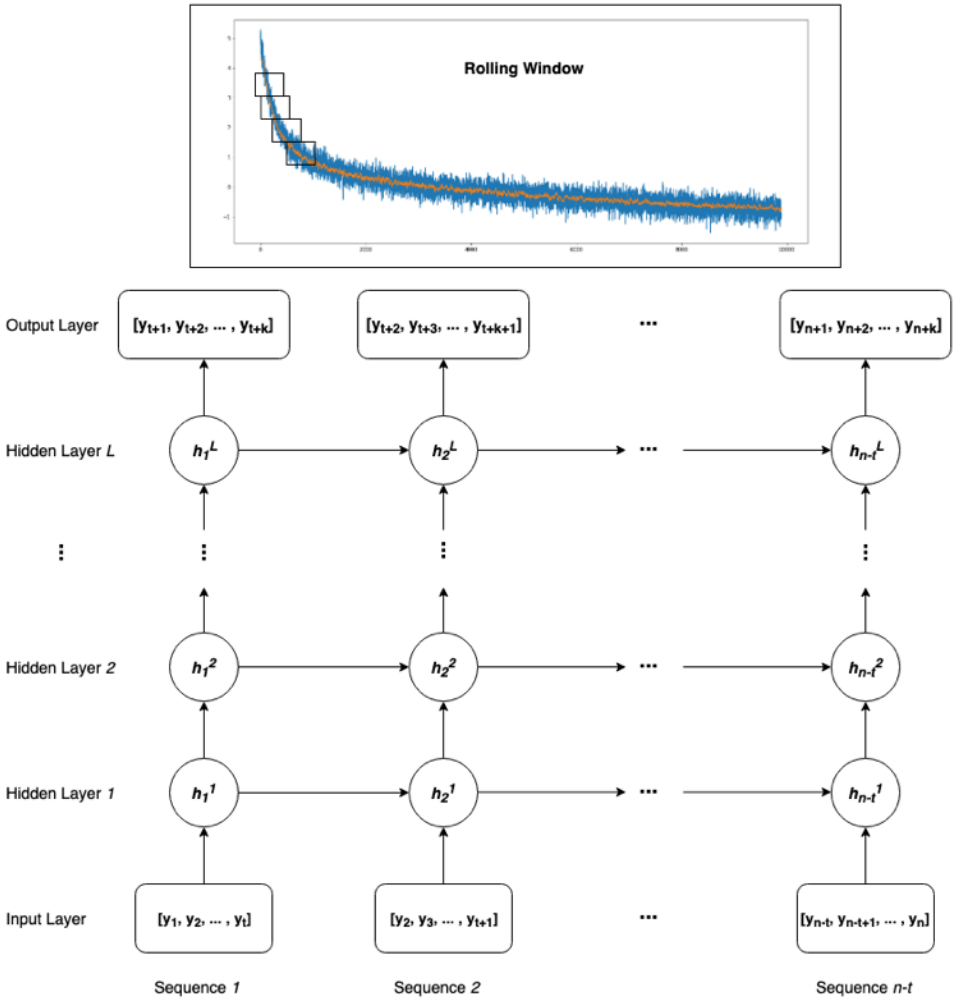


Fig. 17-The structure of the DeLSTM model. A model with L hidden layers and the rolling window length t, k for input layer and output layer respectively. $[n - t]$ denotes the batch sizde of data sequence (Modified from Wu et al. 2017).

Unlike the 3-3-1 (three observations $[y_{t-1}, y_t, y_{t+1}]$ per sequence, three hidden layers $[h_t^1, h_t^2, h_t^3]$, one output $[y_t]$ per sequence) structure in the original paper, we derive a sequence-to-sequence topology to fit the missing imputation problem. For example, if the number of a continual missing period is less than 5 consecutive points, suggested $k \leq 5$ for each output sequence. In practice, to achieve a faster convergence and avoid the gradient exploding while training, $t \geq k$ is preferable. In section 3.3, we will discuss the rolling window method which provides data augmentation and simultaneously increases the accuracy of prediction based on the fact that only one time-series sequence is available.

2.2.3 Savitzky-Golay Cascaded Long Short-Term Memory

We recommend the Savitzky-Golay (SG) polynomial smoothing algorithm as a reliable technique for smoothing time-series data. The fundamental description can be found in the original paper (Savitzky et al. 1964) and a thorough investigation of SG method applied in petroleum signal processing was performed by Reynolds (Reynolds et al. 2005).

The SG method is then coupled with CLSTM and the workflow is illustrated in Fig. 18 below.

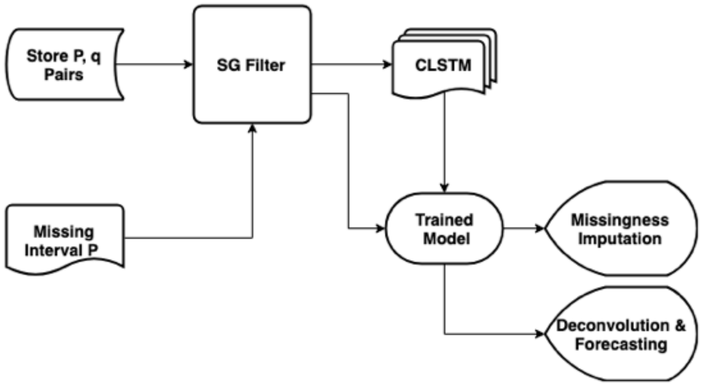


Fig. 18-The workflow of SG-CLSTM for denoising and missingness imputation with both pressure and rate signals presented.

In the SG-CLSTM workflow, both available noisy pressure and rate signals are fed into the SG filter and the resulting denoised signals are then fed into CLSTM for training. The well-trained model is saved after cross-validation. Then, the missing rate interval corresponding filtered pressures are sent to the saved model to reconstruct the missing rates to complete the rate sequence. Finally, the saved model can be used for further deconvolution and forecasting by feeding in an artificial constant-pressure-drop.

2.2.4 Systematic Workflow for Production Surveillance

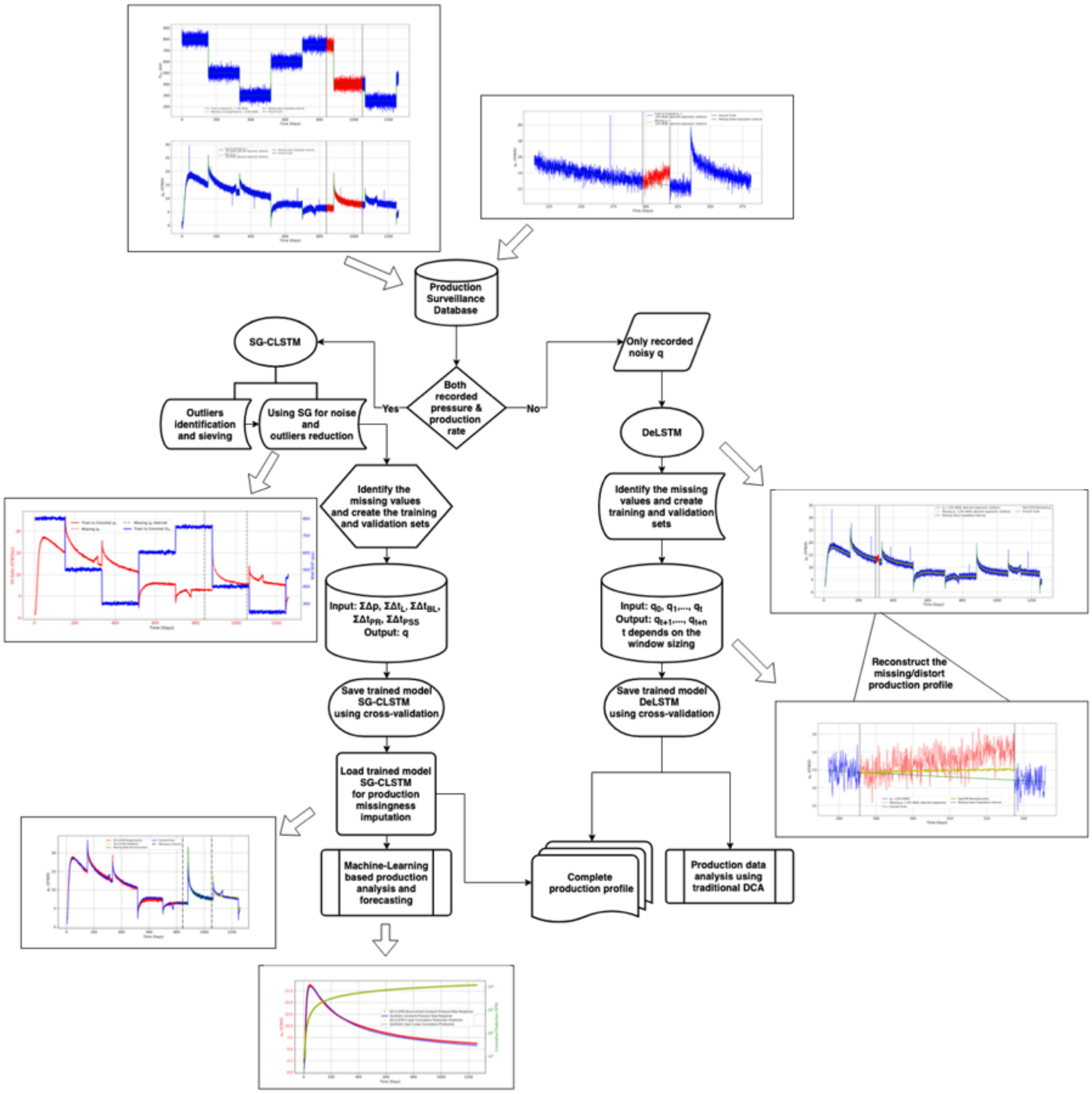


Fig. 19-The systematic workflow combining DeLSTM and SG-CLSTM for production surveillance.

The systematic workflow presented in Fig. 19 illustrates the entire process from production data acquisition to data realization such as production data analysis and forecasting. Two scenarios are considered, one with both pressures and rates recorded. Then in this case, SG-CLSTM will be used for data processing and data realization. In contrast, with only rate signals recorded, DeLSTM will be used only for data processing, however, production data analysis and forecasting are still performed using traditional decline curve analysis (DCA). Both scenarios are demonstrated with one synthetic case and a field case described in Chapter III and IV.

CHAPTER III

APPLICATION TO SYNTHETIC CASE

3.1 Synthetic Case in Rate Transient Analysis

3.1.1 Synthetic Case Model Setup

Machine-learning based algorithms require a certain amount of data for training and validation, and the more data observed by the machine learning model, the more the physics that resides in the data can be captured by the model, allowing more accurate forecasts. However, in the early development of the unconventional reservoirs, the data acquired is quite limited, usually daily data for several months to one or two years. We created two synthetic cases to approximate this situation. The first case is a homogeneous reservoir produced through a horizontal well with a single fracture. We ran the simulation for 11 years with the only first year or first two years being observed by the model as Fig. 20 and Fig. 21 illustrate. The second synthetic case is also a homogeneous reservoir, produced through a horizontal well with multiple transverse fractures. We ran this simulation for 2 years as shown in Fig. 22. The important reservoir parameters for the single-fracture case and the multi-fracture case are summarized in Table 1 and Table 2.

Part of Chapter III is reprinted with permission of Society of Petroleum Engineers from Pan, Y., Zhou, P., Deng, L., Lee, J. 2019. Production Analysis and Forecasting for Unconventional Reservoirs Using Laplacian Echo-State Networks. Proceedings of SPE Western Regional Meeting, San Jose, California, 23-26 April. SPE-195243-MS. DOI:10.2118/195243-MS. Permission conveyed through Copyright Clearance Center, Inc.

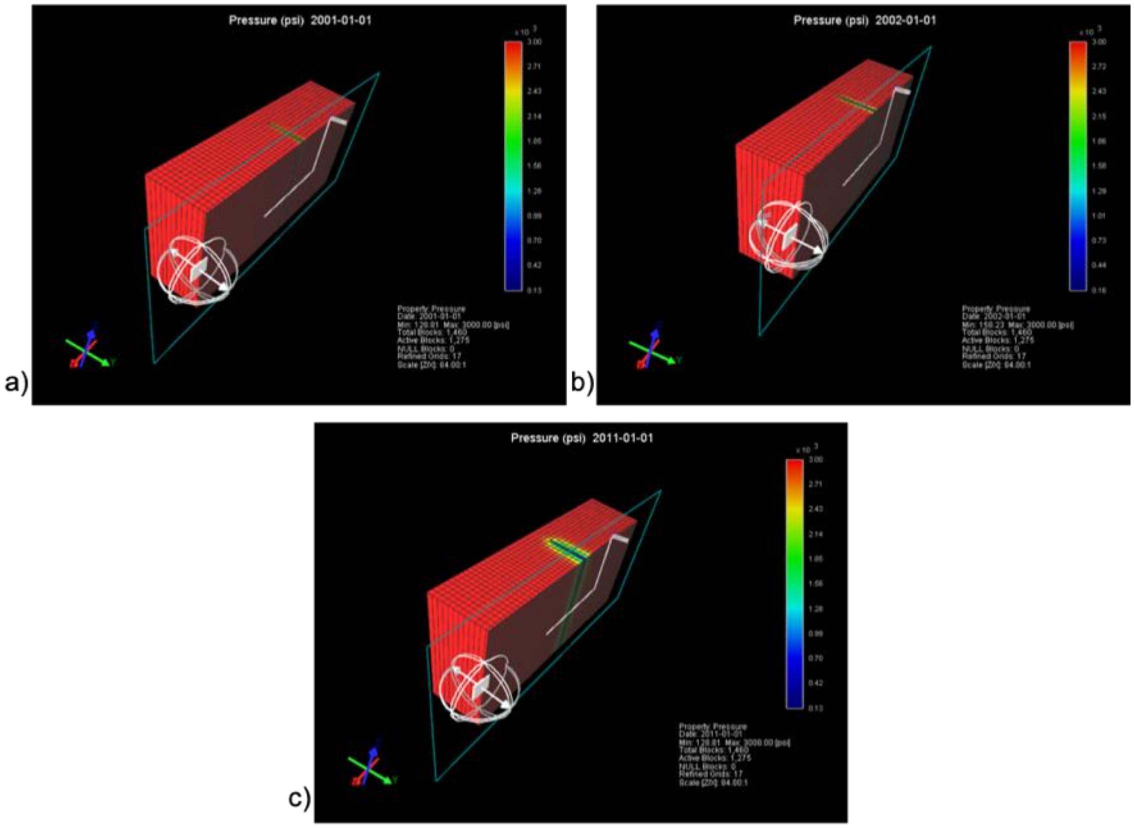


Fig. 20 (a)-Synthetic case reservoir pressure at the end of 1st year; (b)-Synthetic case reservoir pressure at the end of the 2nd year; (c)-Synthetic case reservoir pressure at the end of 11th year.

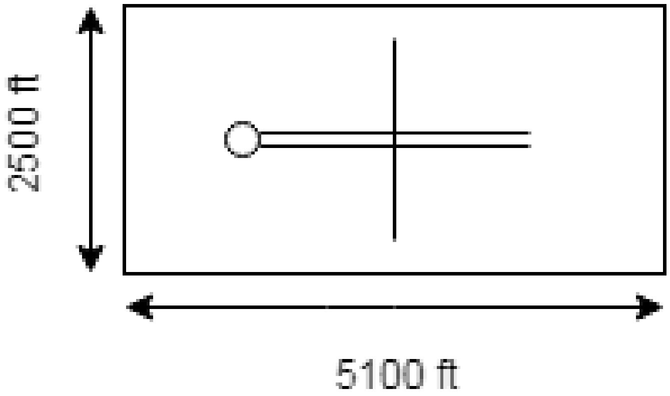


Fig. 21-Top-down view of the reservoir.

k	0.1	md
h	30	ft
ϕ	0.08	
p_i	3000	psi
k_f	5000	md
x_f	800	ft
s_w	0.2	
L_{wb}	1100	ft
μ_o	37	cp
c_t	0.000015	1/psi
Bo	1.028	RB/STB

Table 1-Synthetic Single Fractured Case Simulation Parameters Under Piecewise BHP.

The synthetic test is under piecewise bottom hole pressure (BHP) control so that the pressure drop is not a continuous function of time, and neither is the corresponding rate response. As we stated previously, traditional rate normalized pressure (RNP) or pressure normalized rate (PNR) approaches are not feasible for discontinuous changes in pressure and the corresponding rate response. Four cases will be elucidated in the synthetic single-fractured model: base case, extended base case, noisy data base case, and noisy-data extended base case. After the training procedure is done, the deconvolution test and production forecasting are further performed based on the well-trained proxies (Sun and Ertekin, 2015, 2017).

We then extended the synthetic single-fractured horizontal well case to a synthetic multiple-transverse fracture horizontal well (MTFW) case. This synthetic test is controlled by piecewise BHP in the first year and constant BHP for the next year. The transient flow behavior can be observed in Fig. 22 with the critical reservoir parameters shown in Table 2.

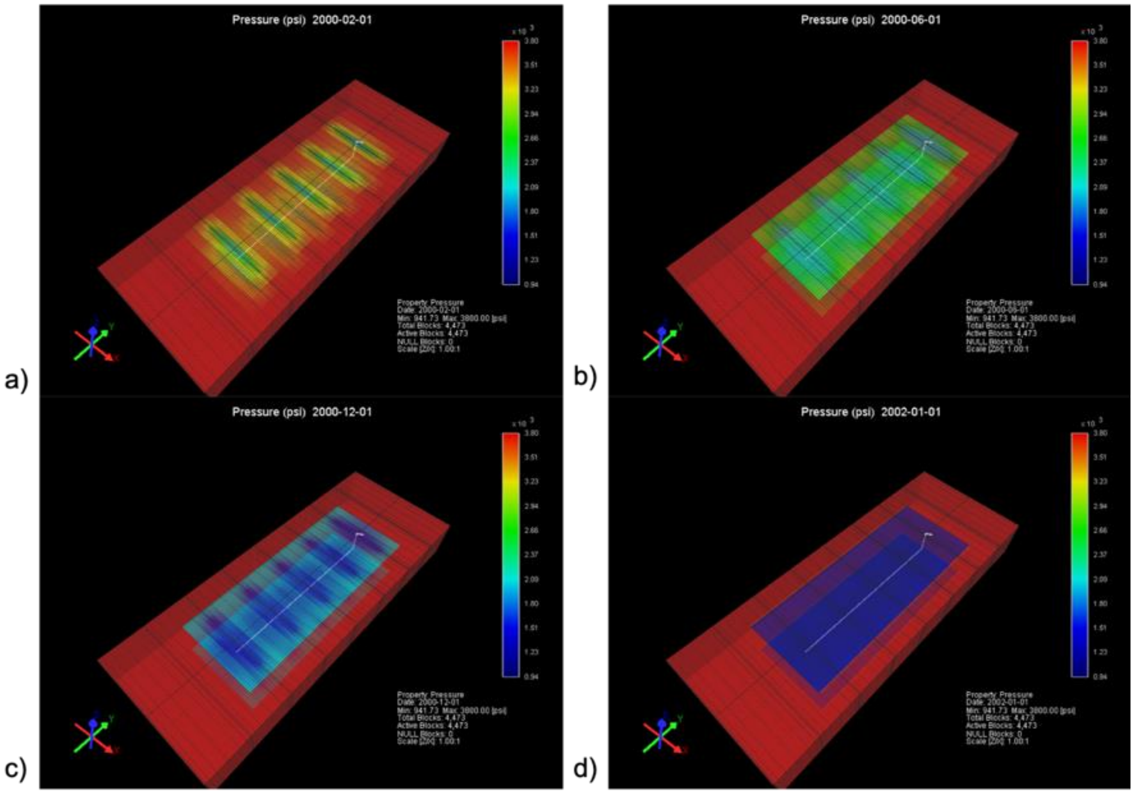
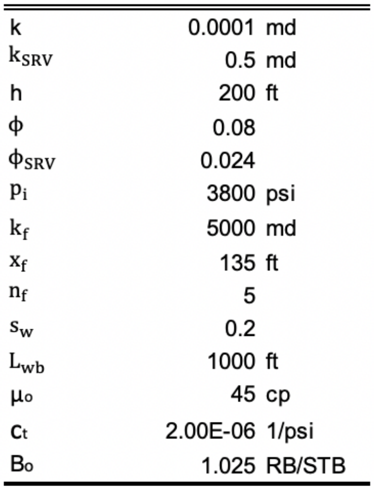


Fig. 22 (a)-Synthetic MTFW case pressure map at the end of 2nd month under piecewise BHP control; (b)-Synthetic MTFW case pressure map at the end of the 6th month under piecewise BHP control; (c)-Synthetic MTFW case pressure map at the end of 12th month under piecewise BHP control; (d)-Synthetic MTFW case pressure map at the end of 24th month under piecewise BHP control.



k	0.0001	md
k_{SRV}	0.5	md
h	200	ft
ϕ	0.08	
ϕ_{SRV}	0.024	
p_i	3800	psi
k_f	5000	md
x_f	135	ft
n_f	5	
s_w	0.2	
L_{wb}	1000	ft
μ_o	45	cp
c_t	2.00E-06	1/psi
B_o	1.025	RB/STB

Table 2-Synthetic MTFW Case Simulation Parameters Under Piecewise BHP.

3.1.2 Single Fracture Base Case

Rate and pressure signals are functions of time and are expressed as $q(t), p(t)$. Both behave chaotically nature in time. Since the primary purpose of this research is to forecast production accurately, we set $q(t)$ as the target. However, the success of training and forecasting is based on constructing physics-based training features. Intuitively, we based training features on superposition in time (Lee et al. 2003) and modified them in addition to the features proposed by Tian and Horne (2019) which are illustrated in Table 3. In the synthetic case, the BHP was changed every two months in the first year as shown in Fig. 23; the corresponding flow rate and model rate regeneration are shown in Fig. 24. Both 1-year signals are exposed to the model for training using Eq. 58 as training inputs and Eq. 59 as training outputs. These physics-based training features are applied thoroughly in the entire study. We first tested LIESN without any additive noise.

Superposition of Pressure Drop	$\Sigma\Delta p(f) = \sum_{j=1}^{\mathrm{N_t}} \left[\Delta p_j - \Delta p_{j-1}\right]$
Superposition with Linear Flow Time	$\Sigma\Delta t_L(f) = \sum_{j=1}^{\mathrm{N_t}} \frac{\left[\Delta p_j - \Delta p_{j-1}\right]}{\left[\sqrt{t_{\mathrm{N_t}} - t_{j-1}}\right]}$
Superposition with Bilinear Flow Time	$\Sigma\Delta t_{BL}(f) = \sum_{j=1}^{\mathrm{N_t}} \frac{\left[\Delta p_j - \Delta p_{j-1}\right]}{\left[\left(t_{\mathrm{N_t}} - t_{j-1}\right)^{\frac{1}{4}}\right]}$
Superposition with Pseudo-Radial Flow Time	$\Sigma\Delta t_{PR}(f) = \sum_{j=1}^{\mathrm{N_t}} \frac{\left[\Delta p_j - \Delta p_{j-1}\right]}{\left[\log\left(t_{\mathrm{N_t}} - t_{j-1}\right)\right]}$
Superposition with Pseudo Steady-State Flow Time	$\Sigma\Delta t_{PSS}(f) = \sum_{j=1}^{\mathrm{N_t}} \frac{\left[\Delta p_j - \Delta p_{j-1}\right]}{\left[t_{\mathrm{N_t}} - t_{j-1}\right]}$

Table 3-Superposition Time Functions for RTA.

Hence, the training features then can be written as below:

$$\mathbf{u_i}(t) = \begin{Bmatrix} \Sigma\Delta p \\ \Sigma\Delta t_L \\ \Sigma\Delta t_{BL} \\ \Sigma\Delta t_{PR} \\ \Sigma\Delta t_{PSS} \end{Bmatrix}, i = 1, \dots, \mathrm{N} \tag{58}$$

and the target $\mathbf{y}(t)$ is simply the production rate array which is a function of t,

$$\mathbf{y}(t) = \{q(t)\} \tag{59}$$

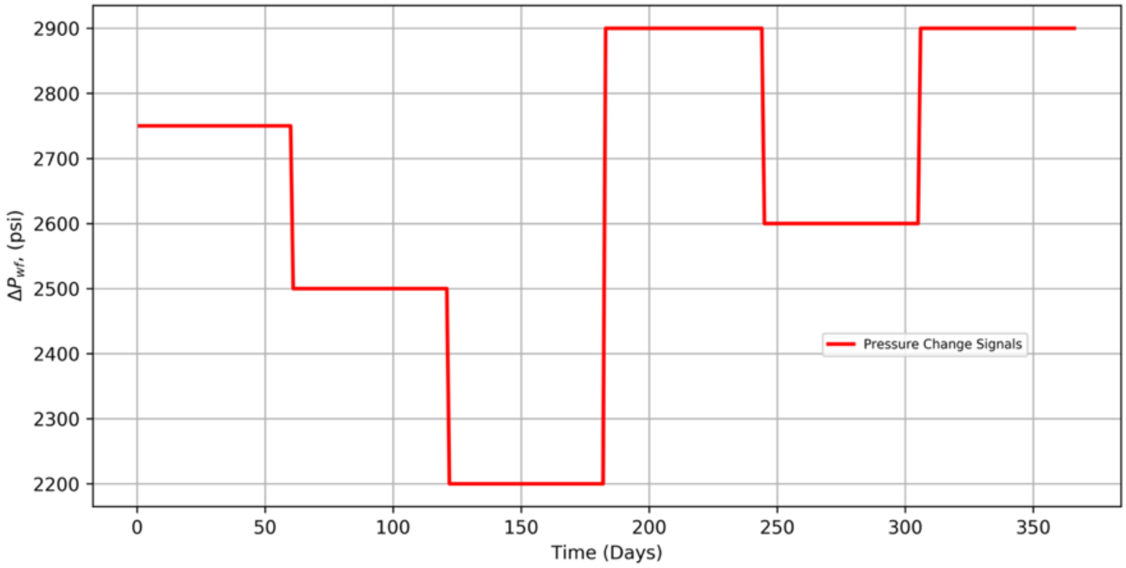


Fig. 23-The pressure changes over time of synthetic base case under piecewise BHP control.

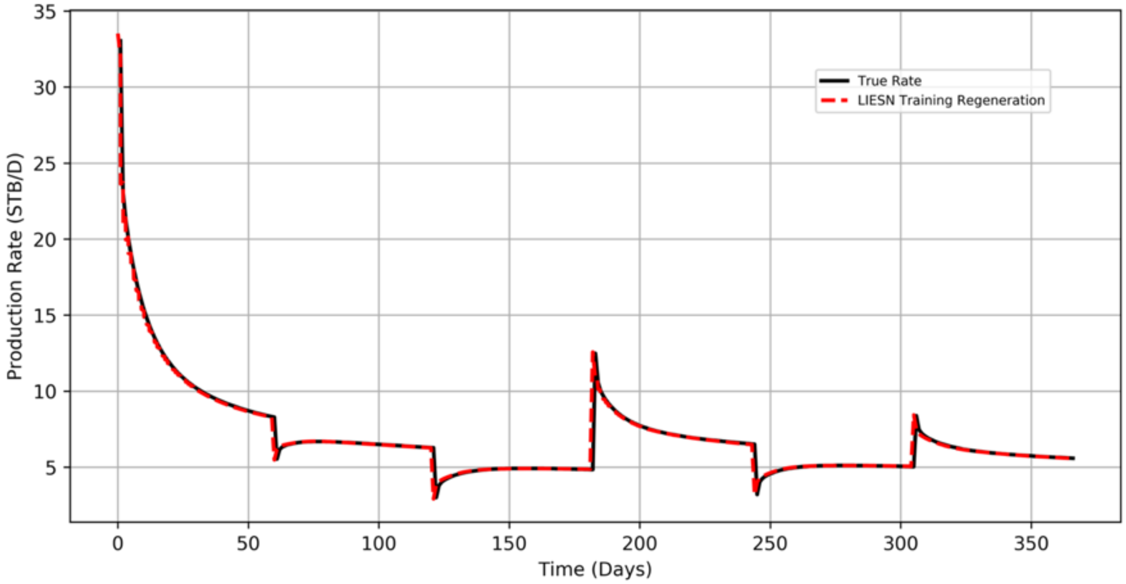


Fig. 24-The corresponding flow rates over time of synthetic base case under piecewise BHP control.

The black solid line in Fig. 24 is the target training output and the red dashed line is the model regeneration after training. We observe a good alignment between the training regeneration and the true rate. This training process has been validated by several trials using cross-validations (CVs). The regularization also avoids the overfitting problem, and using CVs to save the best model is still a required procedure. After the training procedure is finished and the best model is saved, we further perform a deconvolution test and forecast production.

We firstly deconvolve the variable pressure -- variable rate profile into a constant-pressure rate response. By feeding in an artificial constant-pressure-drop (in this case 2750 psi) using physics-based feature transformation from Eq. 58, the rate response is then acquired. Fig. 25 indicates the deconvolution results for the base case.

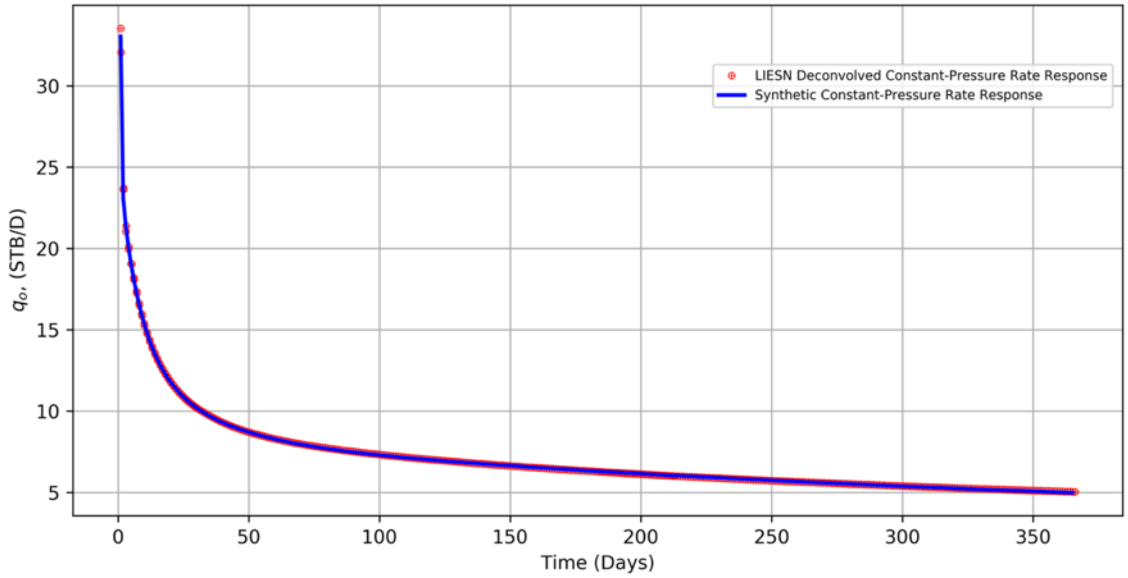


Fig. 25-LIESN deconvolved rate response for the base case comparing with the simulation model true rate response.

In the simulated results, we observed a linear flow regime, so the full deconvolved response compared to simulated constant pressure rate response was based on a $1/q_o$ vs. $\sqrt{t}$ plot as shown in Fig. 26. Again, a good match can be observed; a $1/q_o$ vs. $\sqrt{t}$ plot usually identifies the errors preserved in the deconvolution test. Nonetheless, we see only small deviations at the curve tail where the reservoir being drained under a linear flow regime. We suspect that the training data length is short using 1-year data in which errors are not pervasively aggregated.

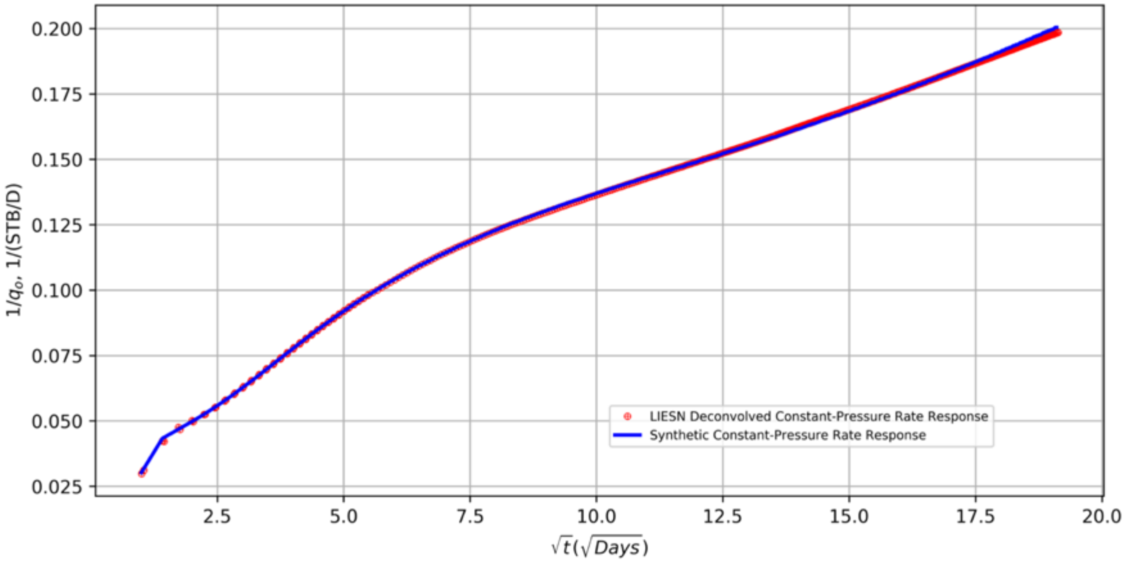


Fig. 26-LIESN deconvolved rate response for the base case compared to the simulation model true rate response on $1/q_o$ $vs.$ $\sqrt{t}$ plot.

The purpose of production data analysis is to identify the flow regimes from the available production data and analytically estimate the reservoir properties such as matrix permeability. To illustrate the procedure, we will show an example of a specialized plot for determining matrix permeability. As Fig. 27 shows, the linear flow regime is extracted. According to Eq. 60 and the fluid properties in Table 1, the reservoir matrix permeability can be calculated through the slope on the plot m=0.0079.

$$k = \left(\frac{31.3B}{mhx_f(p_i - p_{wf})}\right)^2 \frac{\mu}{\phi c_t} \tag{60}$$

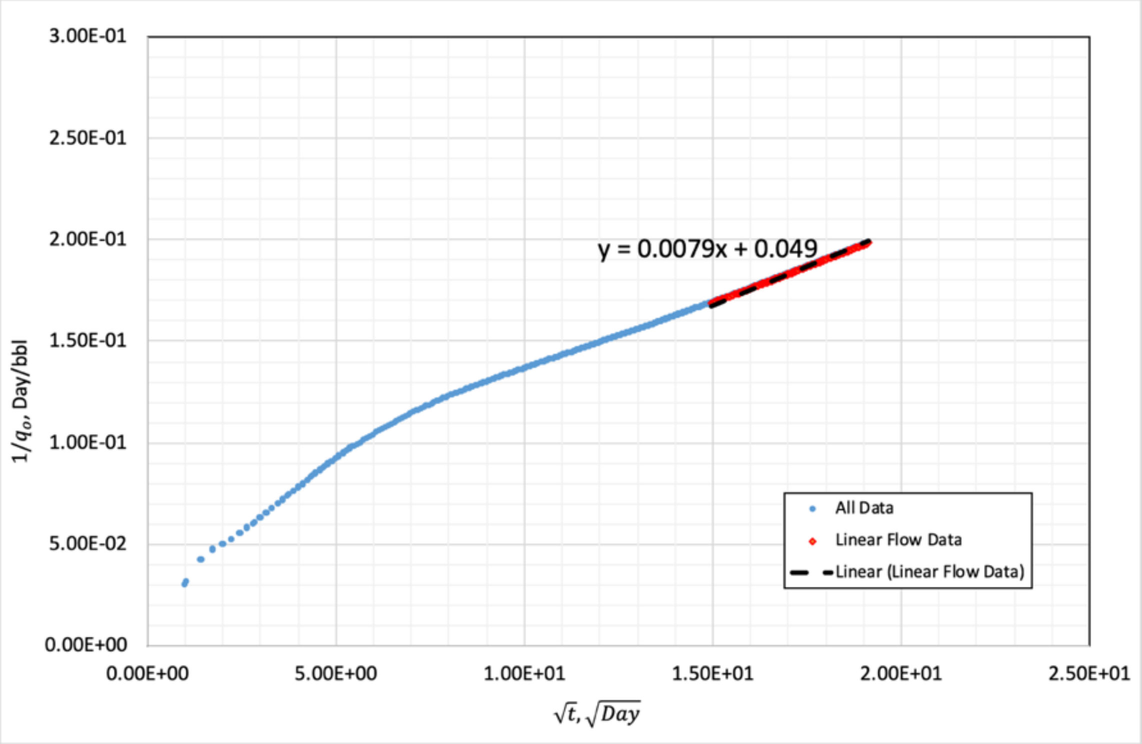


Fig. 27-Base case production analysis using LIESN deconvolved rate response on $1/\mathrm{q_o}\ vs. \sqrt{t}$ specialized plot.

Another purpose is to forecast the production based on the given 1 or 2-year of production histories. The base case used the 1-year based well-trained model to forecast 10 years of future production. The first year's production regeneration combined with the 10-year forecast was then compared with simulation results in Fig. 28.

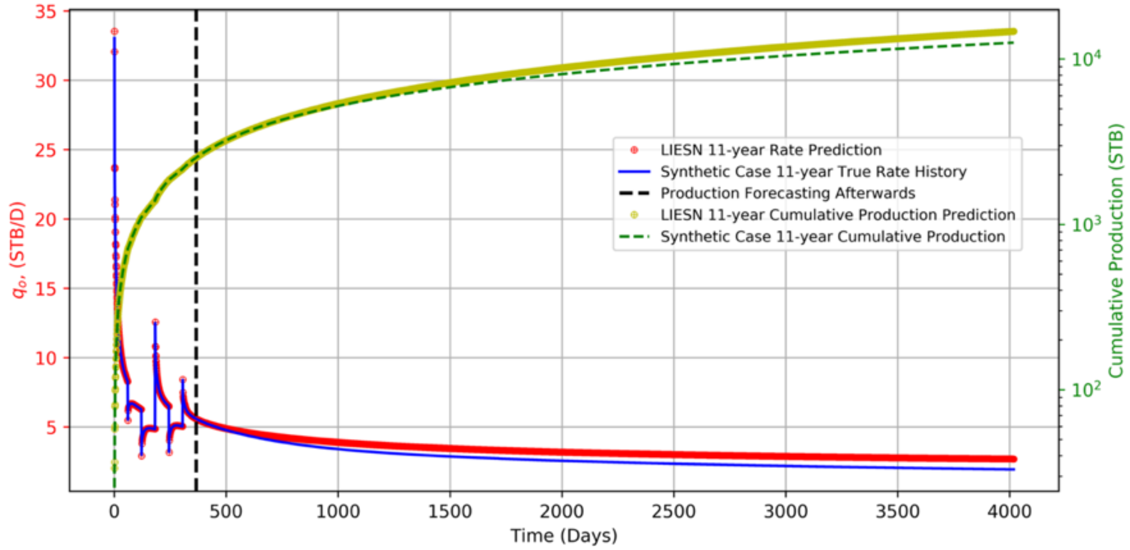


Fig. 28-The comparison of simulated production rate, cumulative production and LIESN forecasted production rate and cumulative production in the base case.

3.1.3 Single Fracture Extended Base Case

With regard to the 1-year training data, the linear flow regime data (roughly $\sqrt{t} > 15$) being exposed to the LIESN model is limited. Therefore, we added a 1-year extension to form a 2-year training data. The purpose of this second test is to further validate the proposed methodology in addition to investigating the potential effect of doubling the training duration. The 2-year BHP, corresponding true rates and model flow rates regeneration are shown in Fig. 29 and Fig. 30.

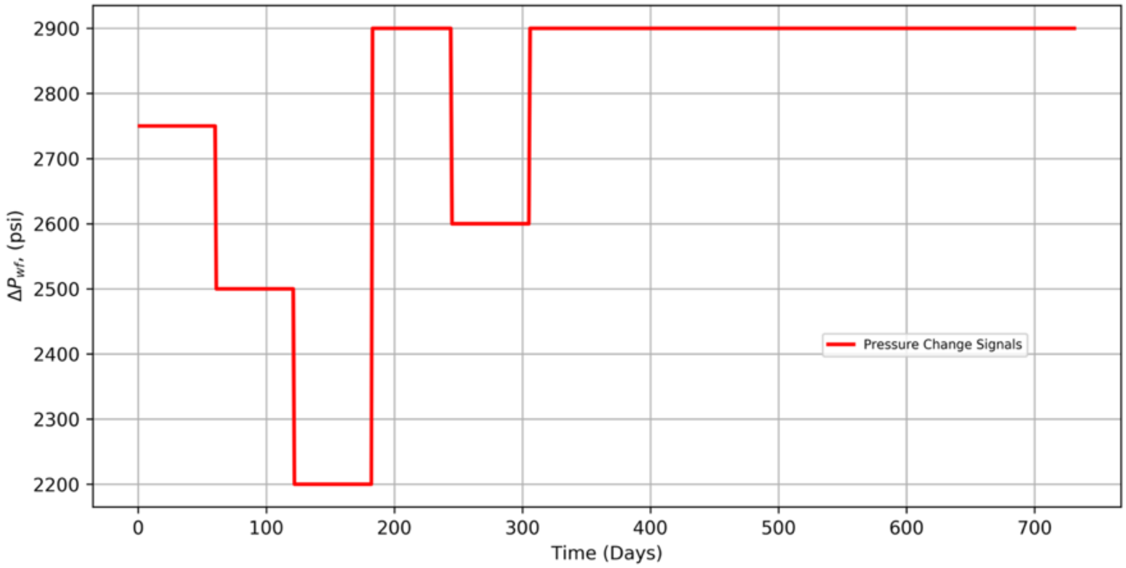


Fig. 29-The pressure changes over time for the extended synthetic base case under piecewise BHP control.

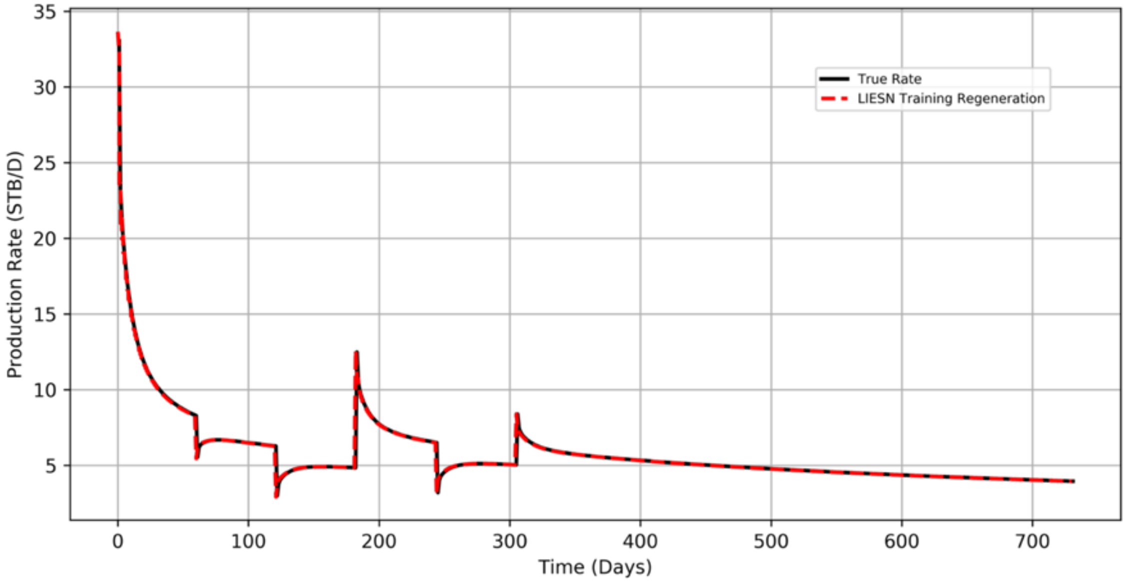


Fig. 30-The corresponding flow rates over time for the synthetic extended base case under piecewise BHP control.

Again, a constant-pressure-drop (2750 psi) was transformed from Eq. 58 and used to determine the rate response. Fig. 31 illustrates the deconvolution results for the extended base case. Similarly, to identify the linear flow regime, the $1/q_o\ vs. \sqrt{t}$ plot in Fig. 32 compares the full deconvolved response and simulated constant pressure rate response.

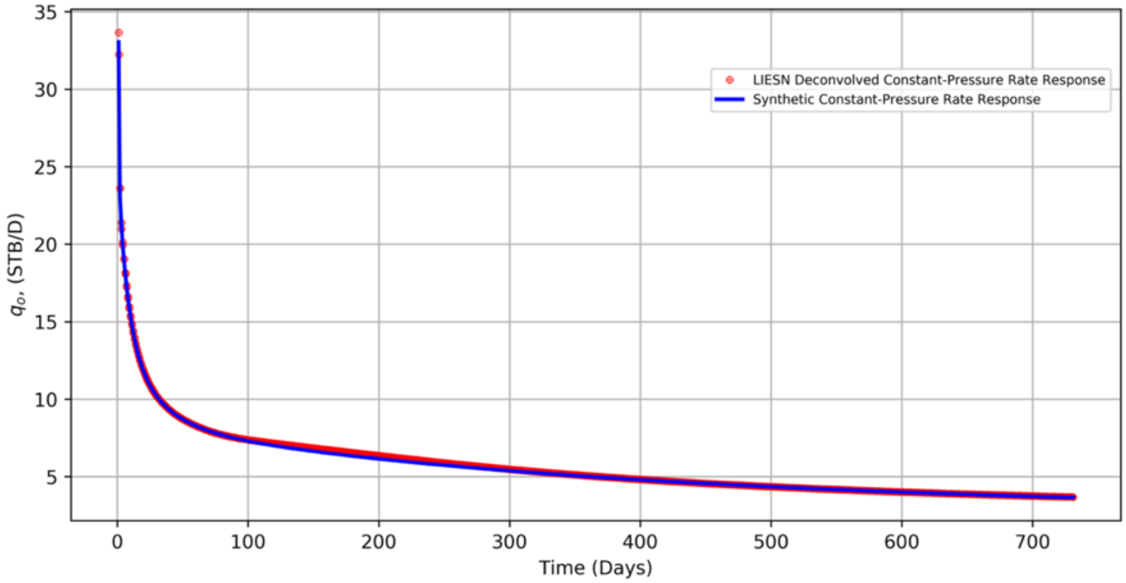


Fig. 31-LIESN deconvolved rate response for the extended base case compared to the simulated rate response.

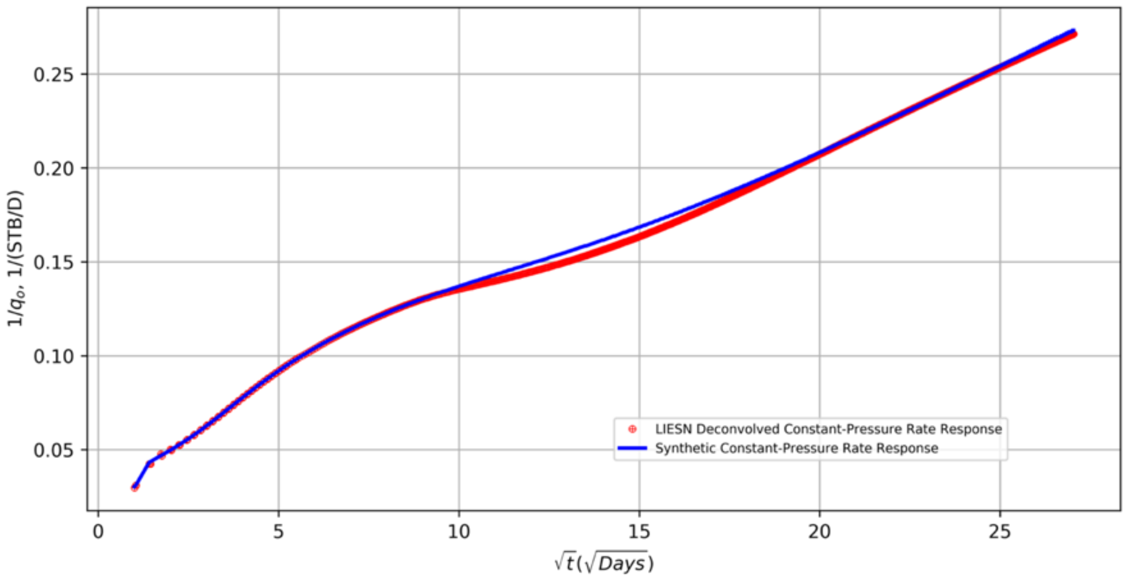


Fig. 32-LIESN deconvolved rate response for the extended base case compared to the simulated rate response.

In Fig. 32, we observe good agreement between true rate and deconvolved rate despite the transition regime. These results align with our purpose of using physics-based training features and training algorithms as key flow regimes are the primary targets to be captured.

The slope of the curve during linear flow was again extracted as shown in Fig. 33. The matrix permeability can then be calculated and compared with other cases during sensitivity analysis.

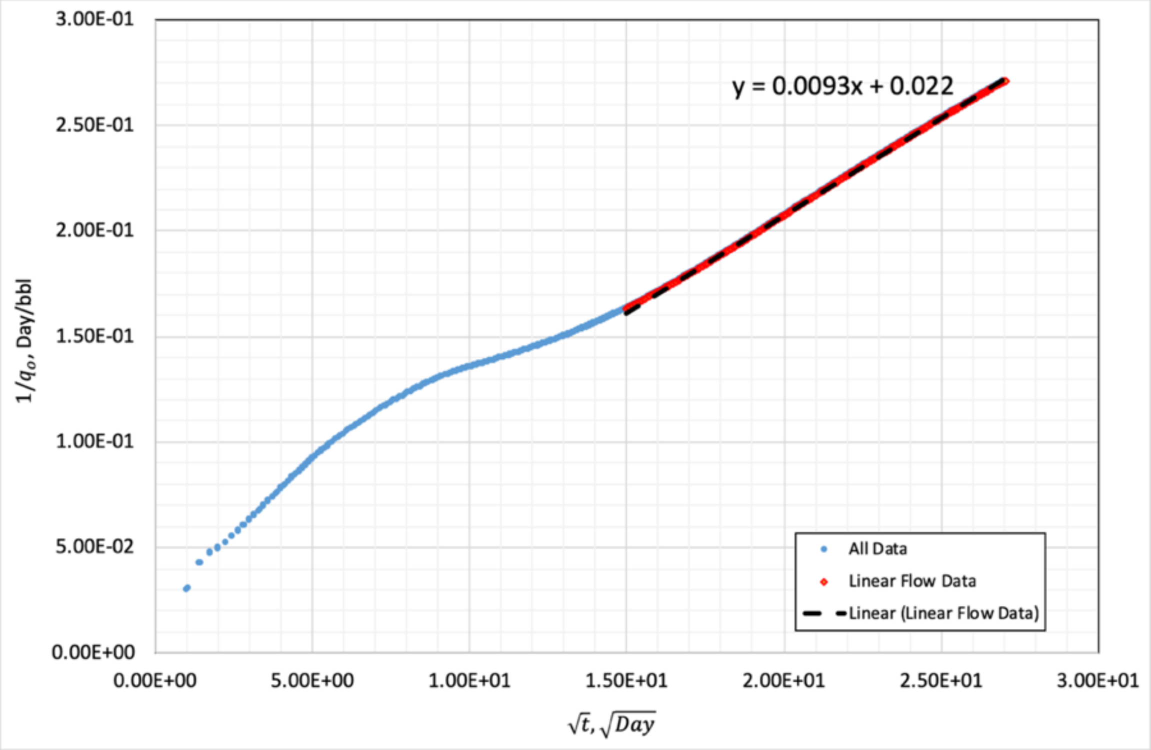


Fig. 33-Extended base case production analysis using LIESN deconvolved rate response on $1/q_o$ $vs.\sqrt{t}$ specialized plot.

The extended base case used a 2-year based well-trained model to forecast 9 years of future production. Again, combining production regeneration for the first two years and the 9-year forecast and comparing with the simulated rate response is shown in Fig. 34.

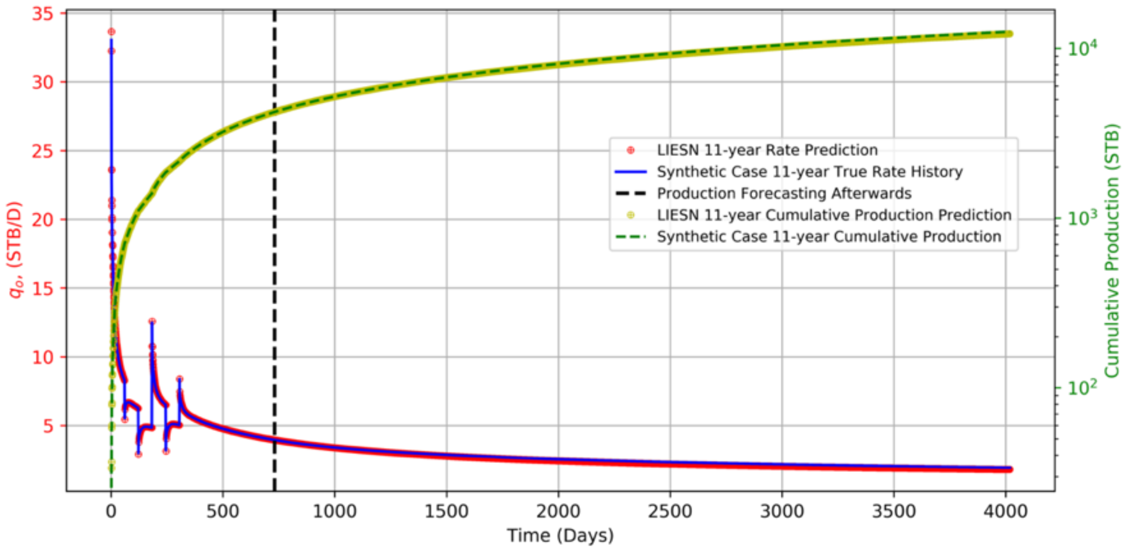


Fig. 34-The comparison of simulated production rates and cumulative production to LIESN forecasted production rates, cumulative production for the extended base case.

3.1.4 Single Fracture Noisy Data Base Case

In actual cases, data gathered downhole including noise could possibly affect deconvolution results and production forecasting. The source of noise may come from system errors or random errors. To study this, we artificially added 5% white Gaussian noise to both BHP and flow rate signals based on the clean data in the base case. The noisy pressure changes, noisy flow rates, and model smoothed rates regeneration are shown in Fig. 35 and Fig. 36. We performed several deconvolution trials using LIESN, and found that the deconvolved rate response was strongly influenced by noise and will not be shown here. Therefore, the LEESN approach was developed and tested to demonstrate its tolerance to noisy data.

The training input features using LEESN are different from those proposed in Eq. 58. We added an additional unit array as shown in Eq. 61. The training output features are the same as Eq. 59.

$$\mathbf{u}_{\mathbf{i}}(t) = \left\{ \begin{matrix} 1 \\ \Sigma\Delta p \\ \Sigma\Delta t_L \\ \Sigma\Delta t_{BL} \\ \Sigma\Delta t_{PR} \\ \Sigma\Delta t_{PSS} \end{matrix} \right\}, i = 1, \dots, \mathrm{N} \tag{61}$$

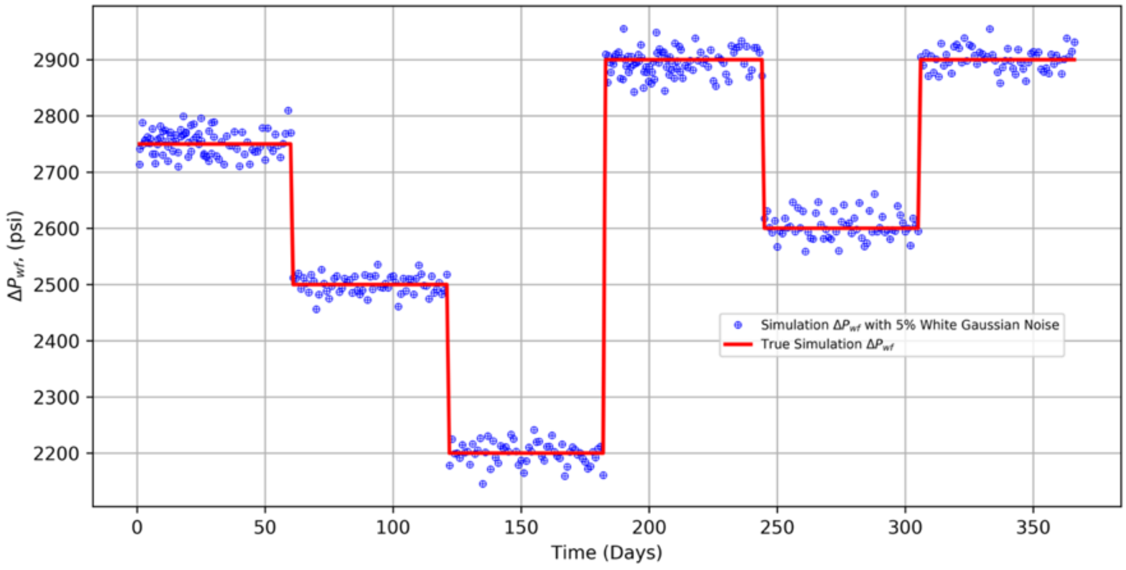


Fig. 35-The true pressure changes and 5% additive white Gaussian noise pressure changes over time for synthetic noisy data base case under piecewise BHP control.

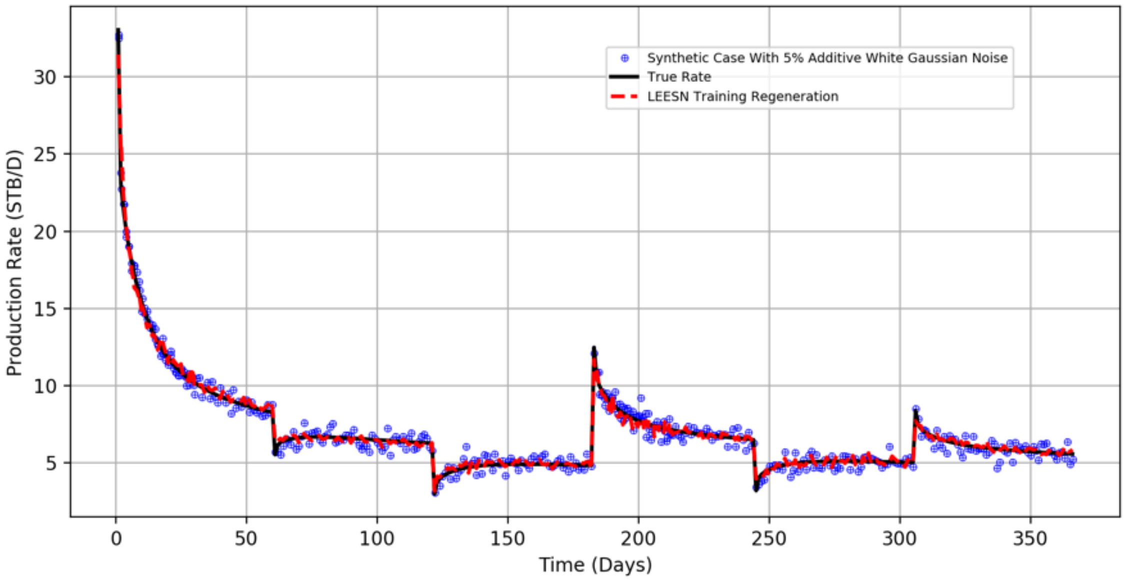


Fig. 36-The true flow rates, 5% additive white Gaussian noise flow rates and model smoothed rates regeneration over time for synthetic noisy data base case under piecewise BHP control.

We continued to use a constant pressure drop (2750 psi) as input into the model to acquire a deconvolved rate response. A constant-pressure rate decline plot and a $1/q_o$ vs. $\sqrt{t}$ specialized plot are presented in Fig. 37 and Fig. 38. From the specialized plot, we can observe that after Laplacian eigenmaps reduced the dimensionality of the original reservoir space, deconvolved linear flow characteristics start to deviate from the true response.

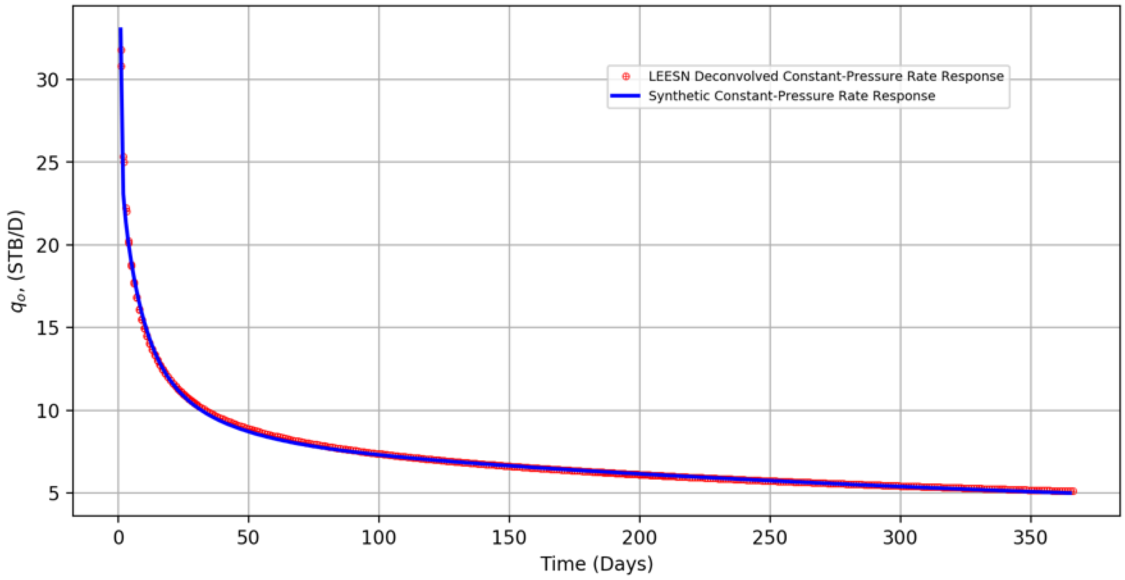


Fig. 37-LEESN deconvolved rate response for synthetic noisy data base case under piecewise BHP control compared to simulated rates.

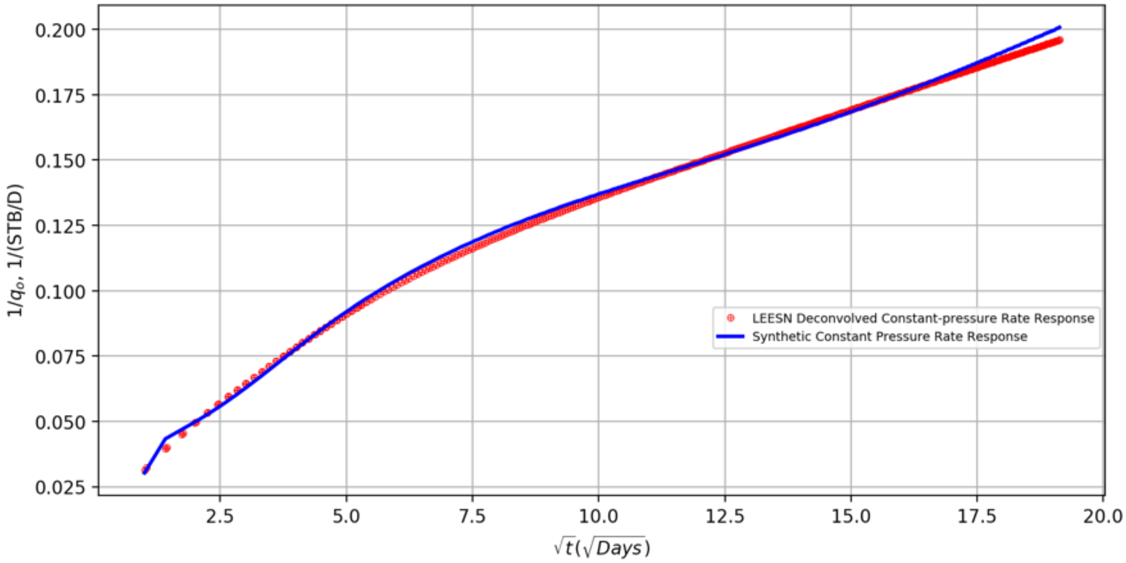


Fig. 38-LEESN deconvolved rate response for synthetic noisy data base case under piecewise BHP control compared to simulated rates.

In Fig. 39 we compute the slope of the linear flow plot from the noisy data base case deconvolution results. From this slope, we can calculate matrix permeability and compare it to the input value in the simulation.

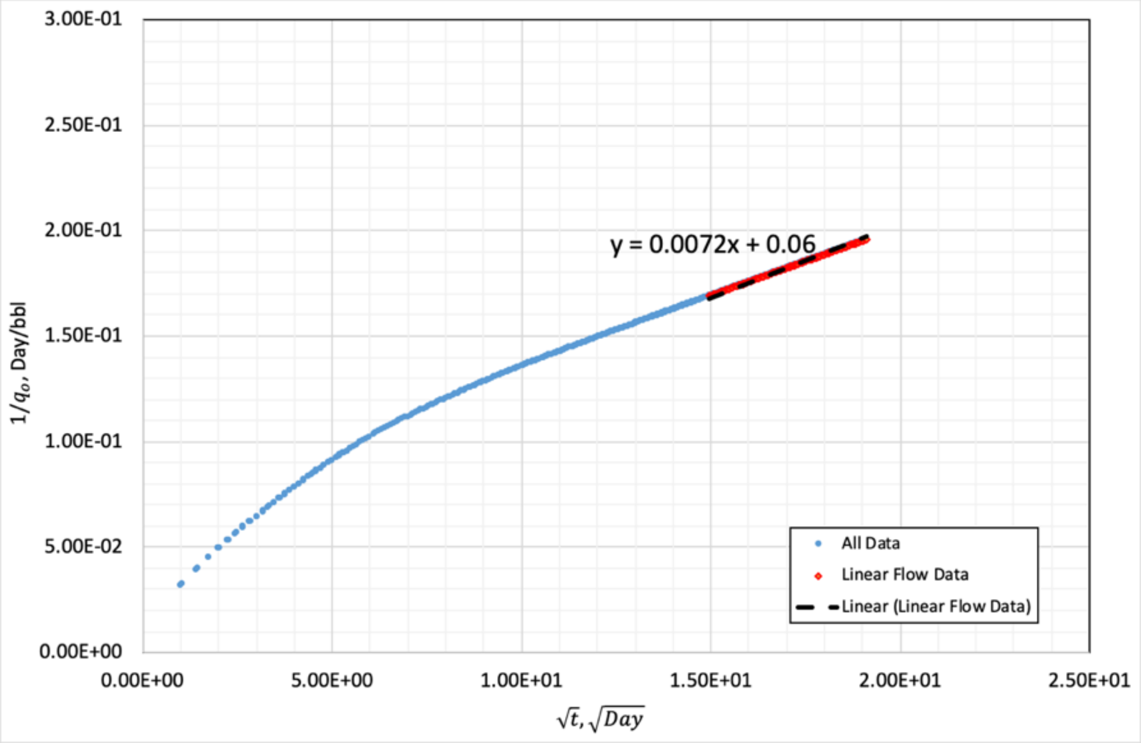


Fig. 39-Noisy data base case production analysis using LEESN deconvolved rate response.

Fig. 40 compares forecasts and simulated production rates and cumulative production.

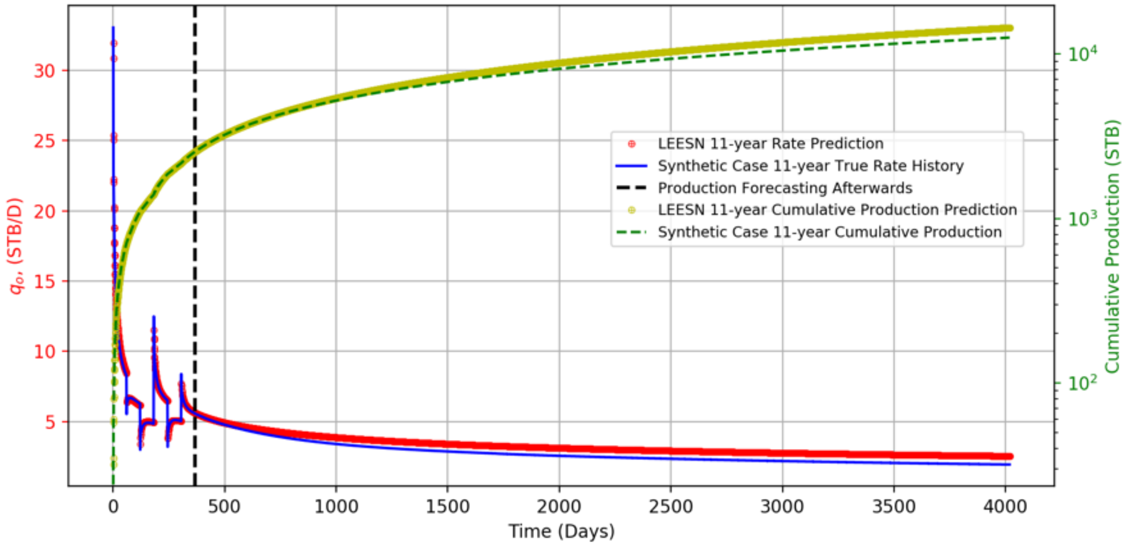


Fig. 40-The comparison of simulated production rate, cumulative production to LEESN forecasted production rate, cumulative production for noisy data base case.

3.1.5 Single Fracture Noisy-Data Extended Base Case

We next added 5% white Gaussian noise to both BHP and flow rate signals based on clean data in the extended base case previously described. Fig. 41 and Fig. 42 present the noisy pressure changes, noisy flow rates and model smoothed rate regeneration separately. We further investigated extended training length using Eqs. 61 and 59 as training input features and output features respectively, and determined how aggregated noise influences on LEESN approach in deconvolution and production forecasting.

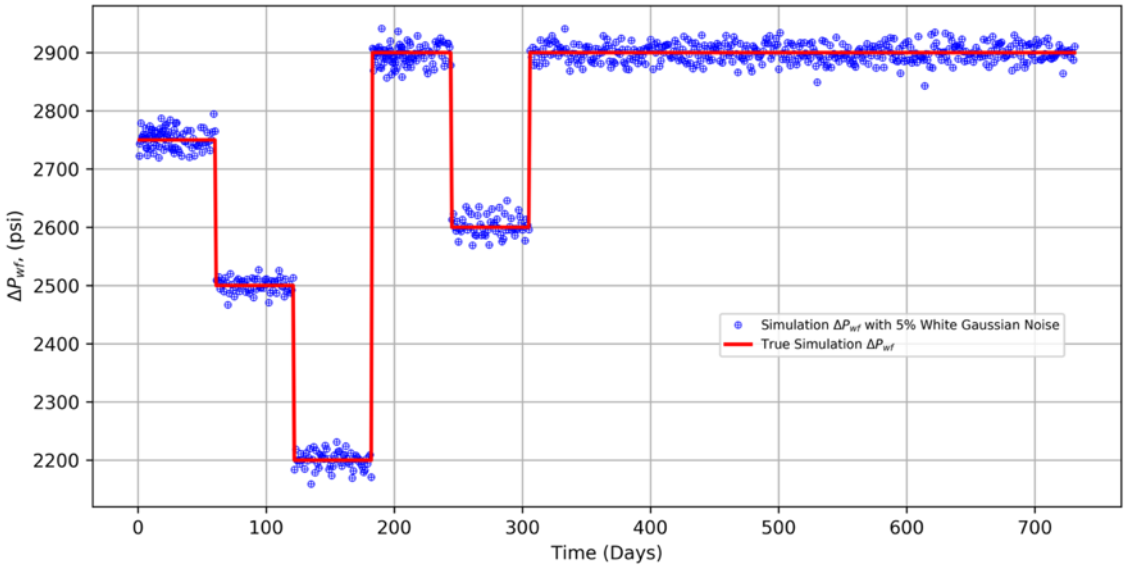


Fig. 41-The simulated pressure changes and pressure chages with 5% additive white Gaussian noise added over time for extended base case under piecewise BHP control.

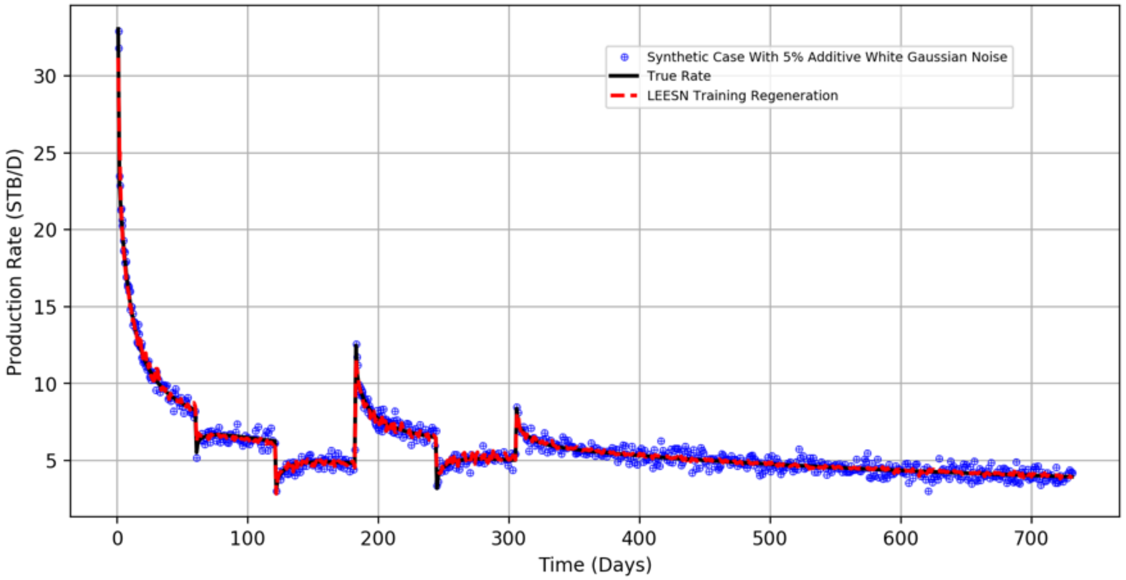


Fig. 42-The simulated flow rates, rates with 5% white Gaussian noise added and model smoothed rates regeneration over time for synthetic noisy extended base case under piecewise BHP control.

The deconvolved rate was generated with the same constant-pressure-drop (2750 psi) and comparing to the simulated rates with constant pressure drop. Fig. 43 and Fig. 44 show the comparison.

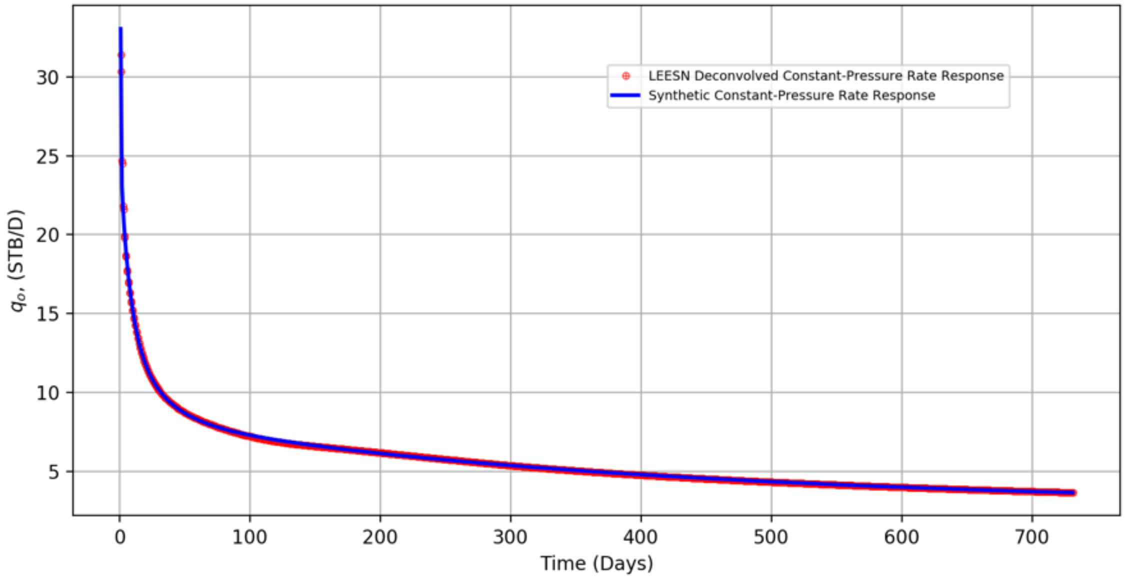


Fig. 43-LEESN deconvolved rate response for the synthetic noisy extended base case under piecewise BHP control compared to the simulated rate response.

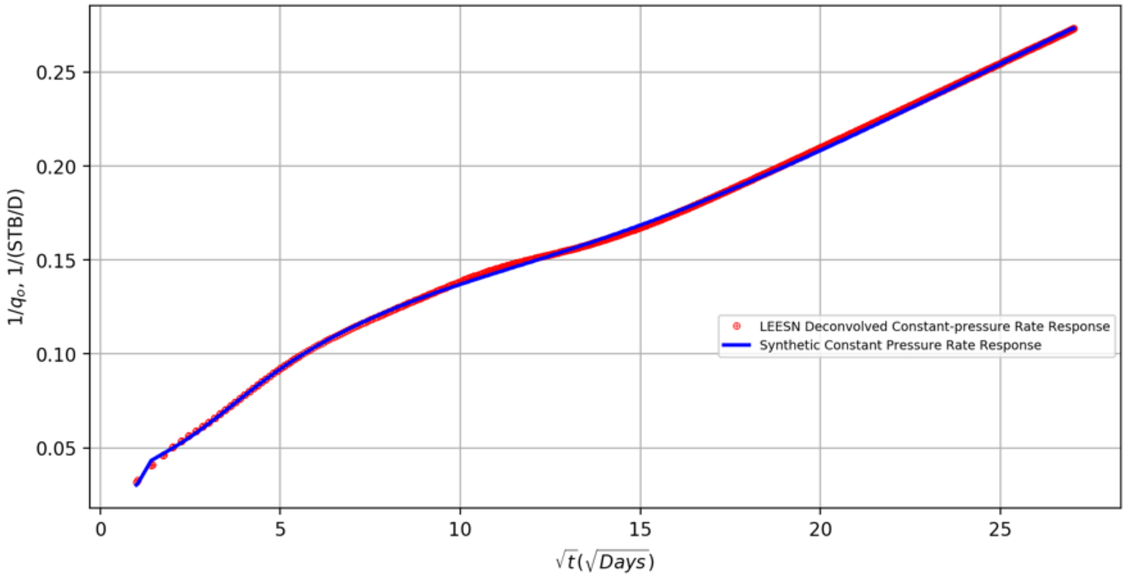


Fig. 44-LEESN deconvolved rate response for the synthetic noisy extended base case under piecewise BHP control compared with simulated rates on $1/q_o$ vs $\sqrt{t}$ plot.

Following the same procedure as before, we generated the specialized plot, Fig. 45, for flow pattern recognition. We estimated the matrix permeability from the slope of the curve in this plot, and we compared it to other cases during sensitivity analysis.

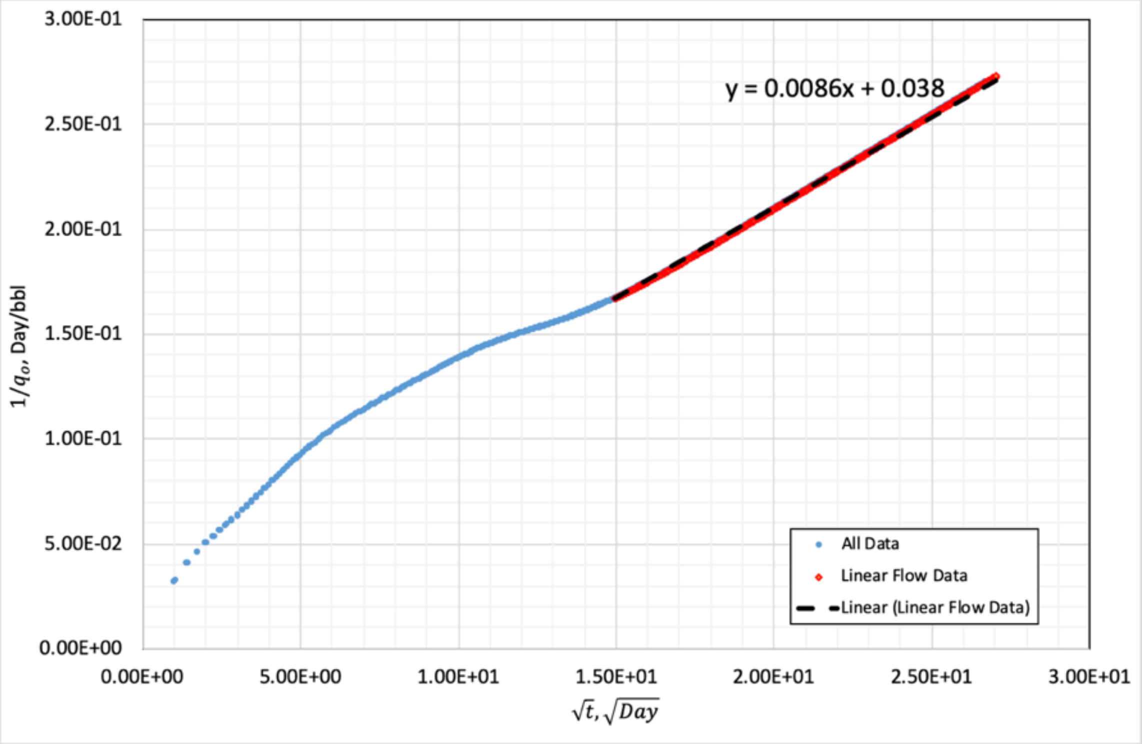


Fig. 45-Noisy-data extended base case production analysis using LEESN deconvolved rate response on $1/q_o$ $vs. \sqrt{t}$ specialized plot.

Similar to the setting of production forecasting in the extended base case, we also forecasted production for the noisy-data extended base case well-trained model for the following 9 years with a combination of curves comparing model forecasts and simulated responses.

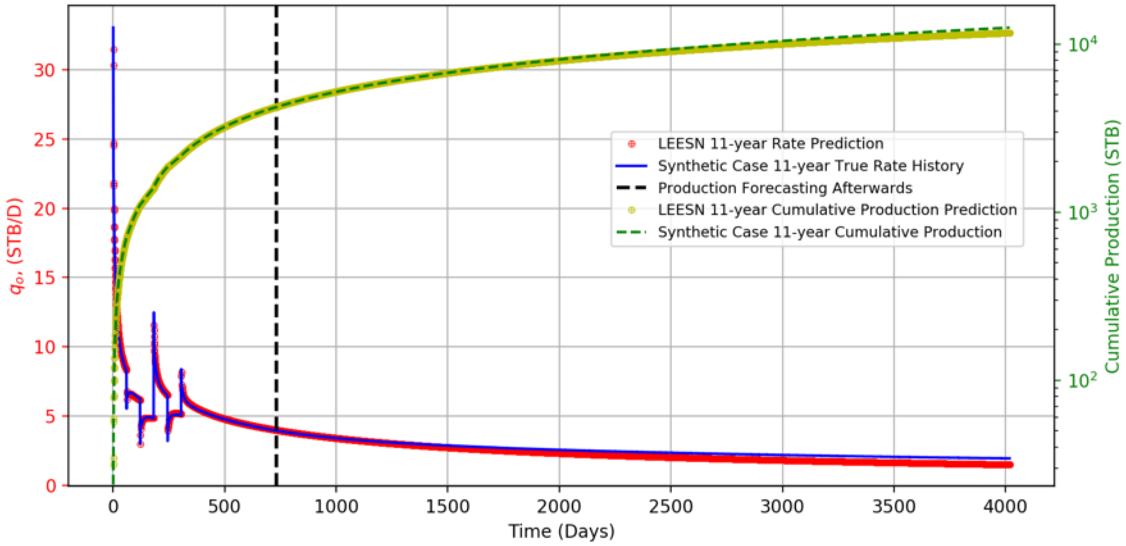
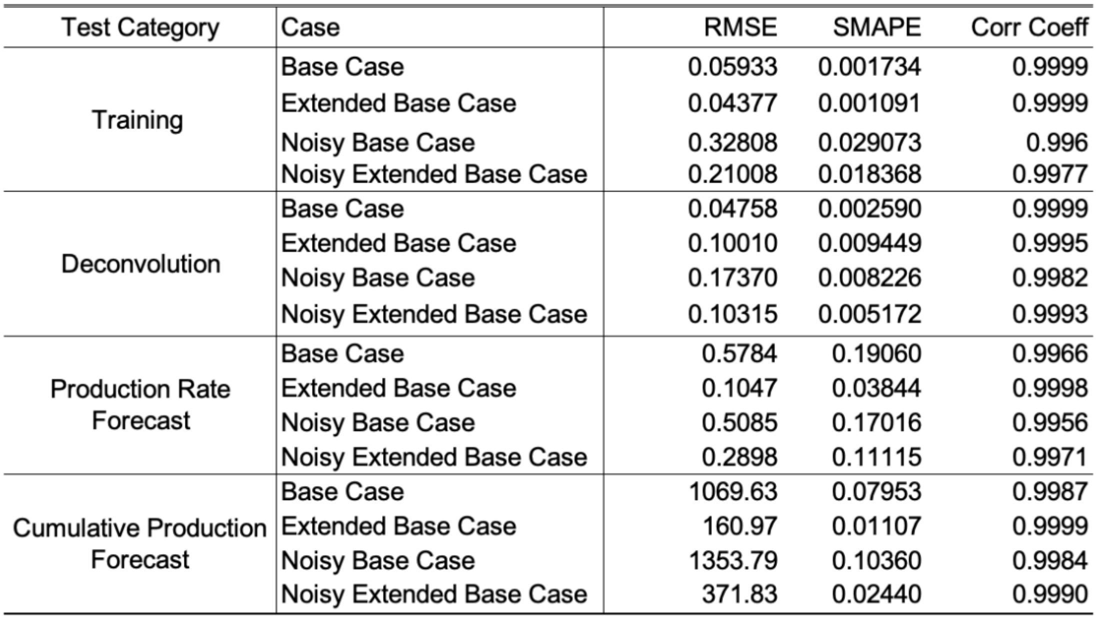


Fig. 46-The comparison between true simulated production rate and cumulative production and LEESN forecasted production rate and cumulative production for the noisy-data extended base case.

3.1.6 Single Fracture Case Sensitivity Analysis

The impact of noise on deconvolution in production data analysis and in production forecasting has not been explored thoroughly in a quantitative manner. We summarize errors in Table 4 by comparing four categories of tests among various synthetic cases. Root mean square error (RMSE), symmetric mean absolute percentage error (SMAPE) and correlation coefficient are used to evaluate the results from different cases. In the training regeneration process results as shown in Fig. 36 and Fig. 42 and training errors in Table 4, LEESN is able to smooth the noisy data and approximate the synthetic true rate. Table 4 also indicates that LISEN applied to extended base case outperforms other as approaches because the data are clean and the training length was increased. In the cases with noisy data, errors in deconvolution and production forecasting are still acceptable. Another intriguing trend is that LEESN with extended data duration performs better both in forecasting and in deconvolution. The deviation during the linear flow regime caused by noise may lead to a poorer representation of the reservoir model. This phenomenon conforms with our expectation as Laplacian eigenmaps which reduce the noise perturbations could also obscure the characteristics of the original data. Nonetheless, the extended training data duration is more critical as it promotes more confidence in identifying flow patterns.

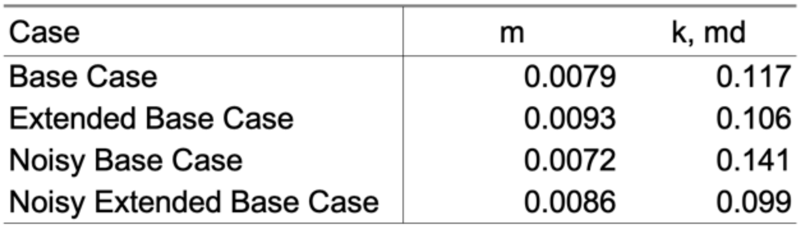


Test Category	Case	RMSE	SMAPE	Corr Coeff
Training	Base Case	0.05933	0.001734	0.9999
	Extended Base Case	0.04377	0.001091	0.9999
	Noisy Base Case	0.32808	0.029073	0.996
	Noisy Extended Base Case	0.21008	0.018368	0.9977
Deconvolution	Base Case	0.04758	0.002590	0.9999
	Extended Base Case	0.10010	0.009449	0.9995
	Noisy Base Case	0.17370	0.008226	0.9982
	Noisy Extended Base Case	0.10315	0.005172	0.9993
Production Rate Forecast	Base Case	0.5784	0.19060	0.9966
	Extended Base Case	0.1047	0.03844	0.9998
	Noisy Base Case	0.5085	0.17016	0.9956
	Noisy Extended Base Case	0.2898	0.11115	0.9971
Cumulative Production Forecast	Base Case	1069.63	0.07953	0.9987
	Extended Base Case	160.97	0.01107	0.9999
	Noisy Base Case	1353.79	0.10360	0.9984
	Noisy Extended Base Case	371.83	0.02440	0.9990

Table 4-Case Error Performance in Various Tests.

From specialized plot Fig. 27, Fig. 33, Fig. 39 and Fig. 45, the slope and calculated permeability are summarized in

. These results should be conformed with a permeability of 0.1 md used to simulate the “clean” production profiles. Errors present in the training data can possibly cause the deconvolved response to deteriorate. However, a good approximation can still be acquired using LEESN with noisy data. These results further demonstrate the applicability of LIESN and LEESN.



Case	m	k, md
Base Case	0.0079	0.117
Extended Base Case	0.0093	0.106
Noisy Base Case	0.0072	0.141
Noisy Extended Base Case	0.0086	0.099

Table 5-Permeability Calculation in Different Cases.

3.1.7 Single Fracture Analysis Using BESN

Section 2.1.4 presented an alternative linear solver, Bayesian Ridge Regression (BRR), as an alternative to Ridge Regression. In this part of the study, we used the same synthetic single-fracture noisy data base case in Section 3.1.4 to validate the regression performance of BRR. Fig. 47 and Fig. 48 illustrate the deconvolution test case we analyzed. Similar deconvolution results can be achieved with LEESN.

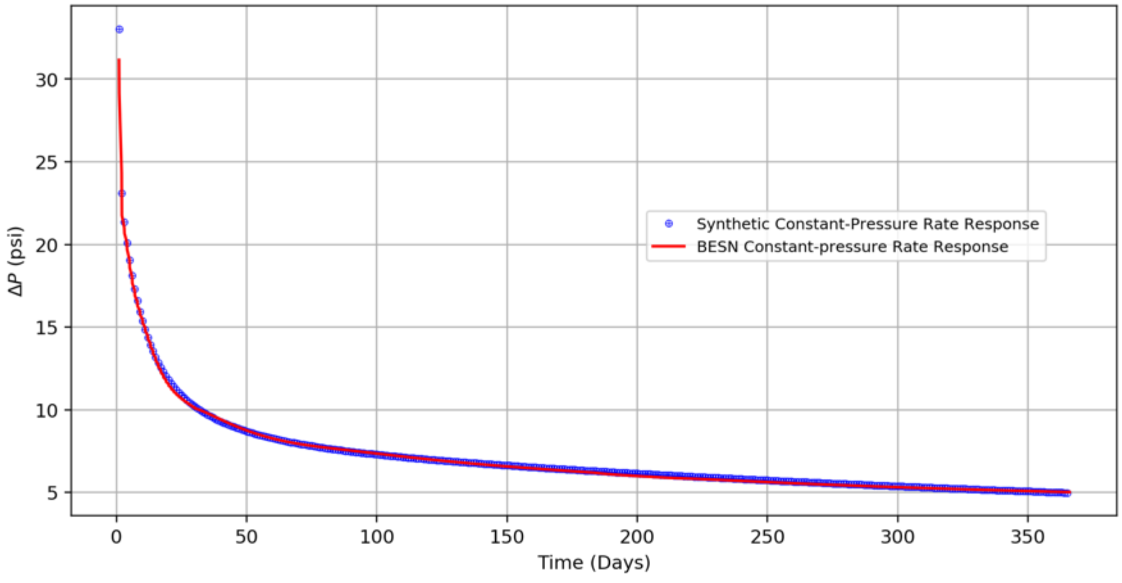


Fig. 47-BESN deconvolved rate response for synthetic noisy data base case under piecewise BHP control compared to simulated rates.

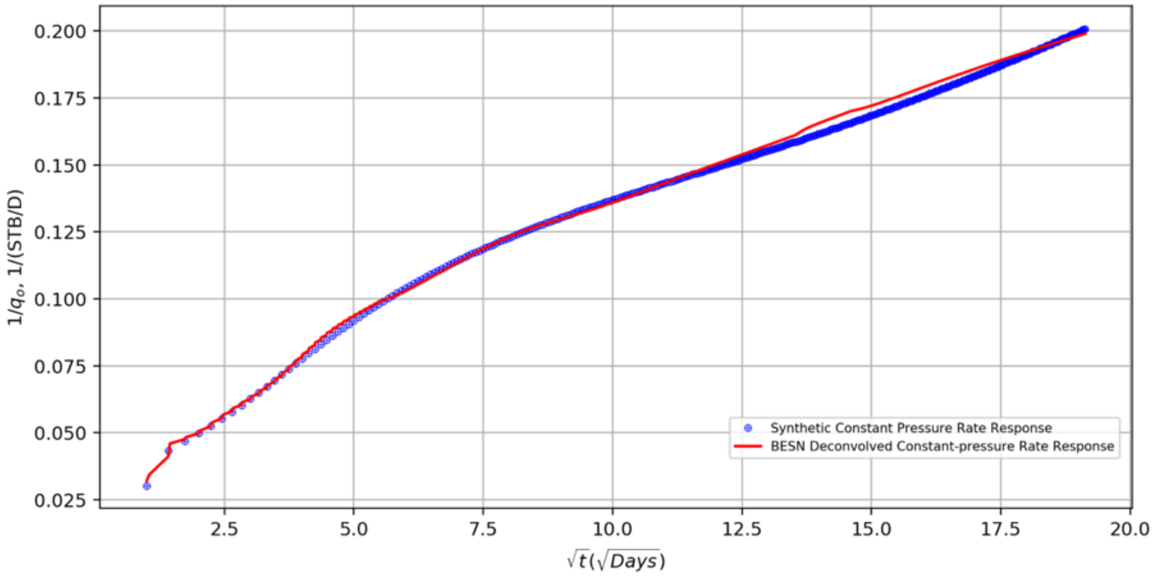


Fig. 48-BESN deconvolved rate response for synthetic noisy data base case under piecewise BHP control compared with simulated rates on $1/q_o$ vs $\sqrt{t}$ plot.

The value of BESN is that it provides error intervals, assuming Gaussian distribution of noise. Using BESN, we forecasted 10 years of production rates and cumulative production and compared with simulated values. Fig. 49 illustrates 10% standard deviation intervals in addition to the mean prediction. This confidence interval would highly likely to cover most future production outcomes.

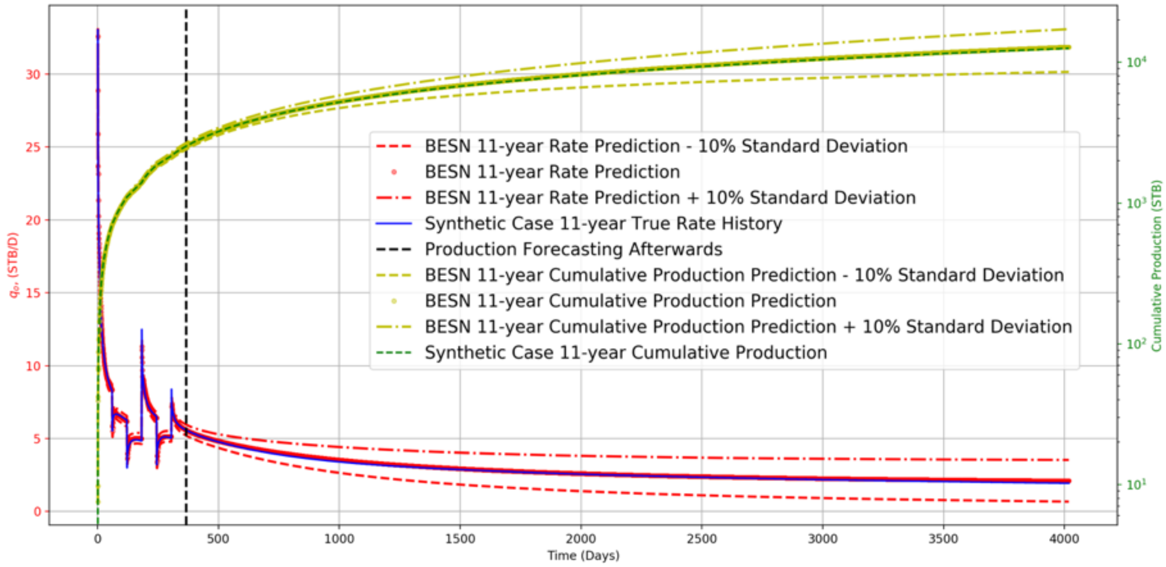


Fig. 49-The comparison of simulated production rates, cumulative production and BESN forecasted production rate with 10% error interval, cumulative production with 10% error interval on noisy data base case.

3.1.8 Multiple Transverse Fractures Noisy Case

In section 3.1.1 we described a case in which we simulated synthetic production data for a horizontal well with multiple transverse hydraulic fractures (MTFW). We described the variable operation control with piecewise bottom-hole pressures changing every two months during a one-year simulation. The pressure drawdown for the entire year is illustrated in Fig. 50. We employed the LEESN smoothing procedure using Eq. 61 and Eq. 59 as the training features. The training regeneration results are shown in Fig. 51.

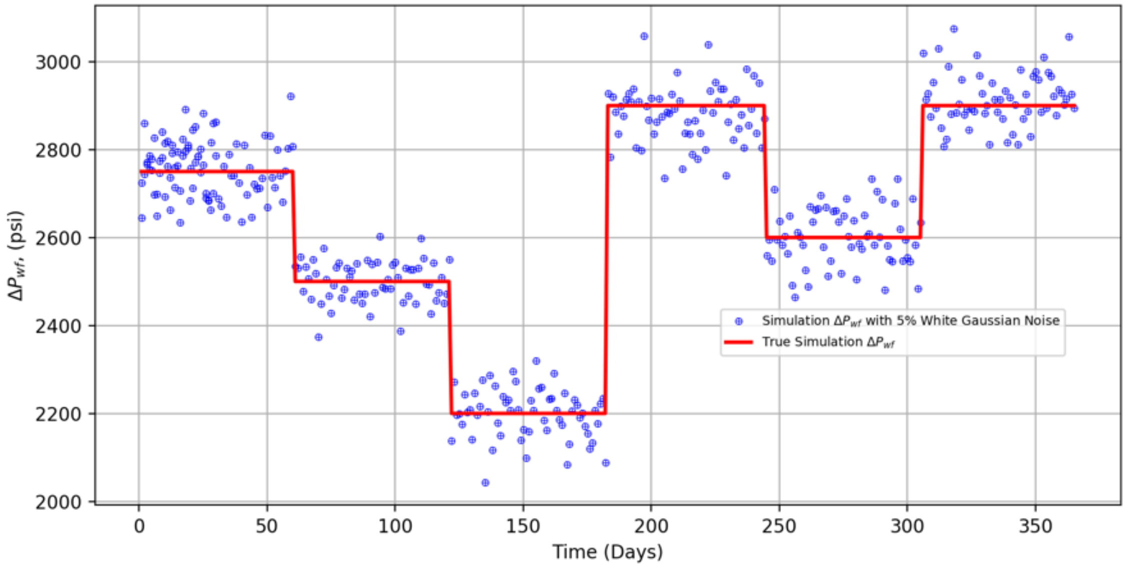


Fig. 50-The simulated pressure changes and pressure changes with 5% white Gaussian noise added over time for synthetic noisy-data MTFW case under piecewise BHP control.

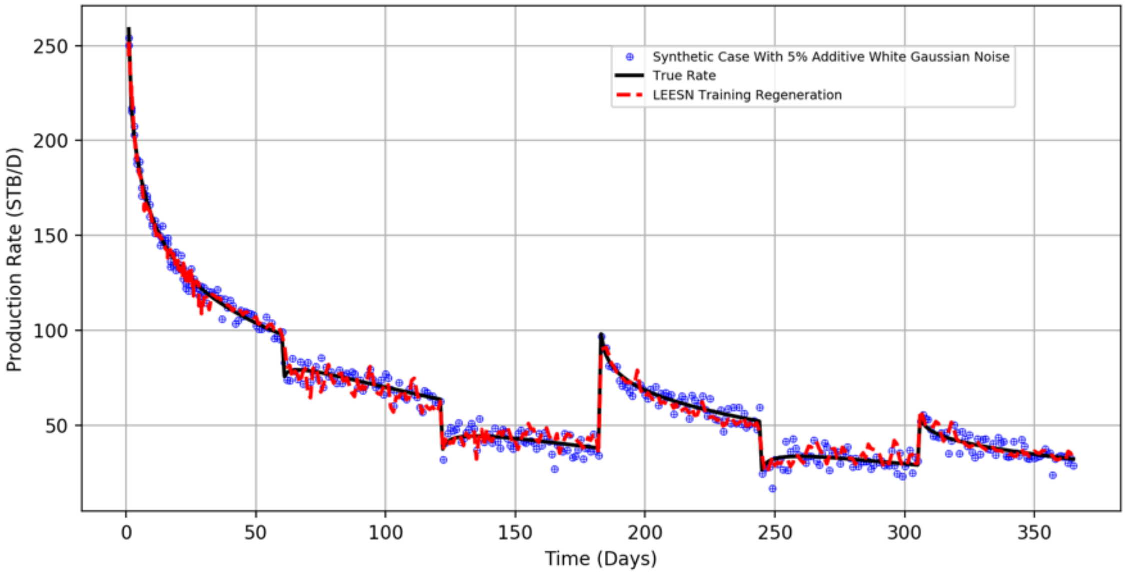


Fig. 51-Simulated flow rates, noisy-data rates with 5% added white Gaussian noise and model-smoothed rate regeneration over time for synthetic noisy-data MTFW case under piecewise BHP control.

Similar to the single-fracture cases, the constant-pressure-drawdown was set at 2750 psi. We compared the deconvolution performance to the simulation results shown in both a cartesian plot, Fig. 52, and a specialized plot, Fig. 53.

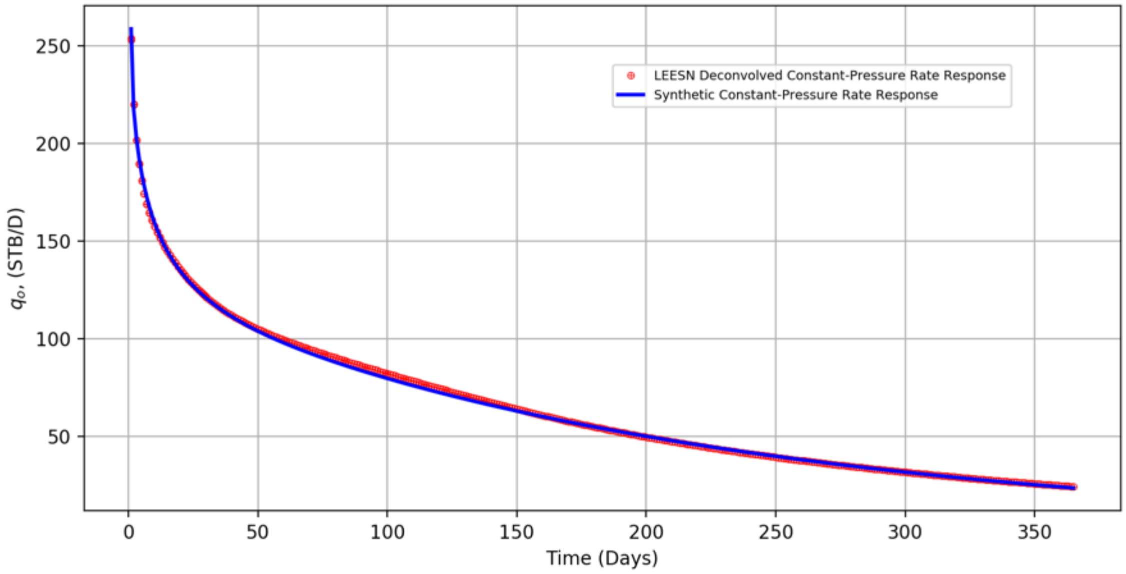


Fig. 52-LEESN deconvolved rate response for the synthetic MTFW case under piecewise BHP control compared to simulated rates.

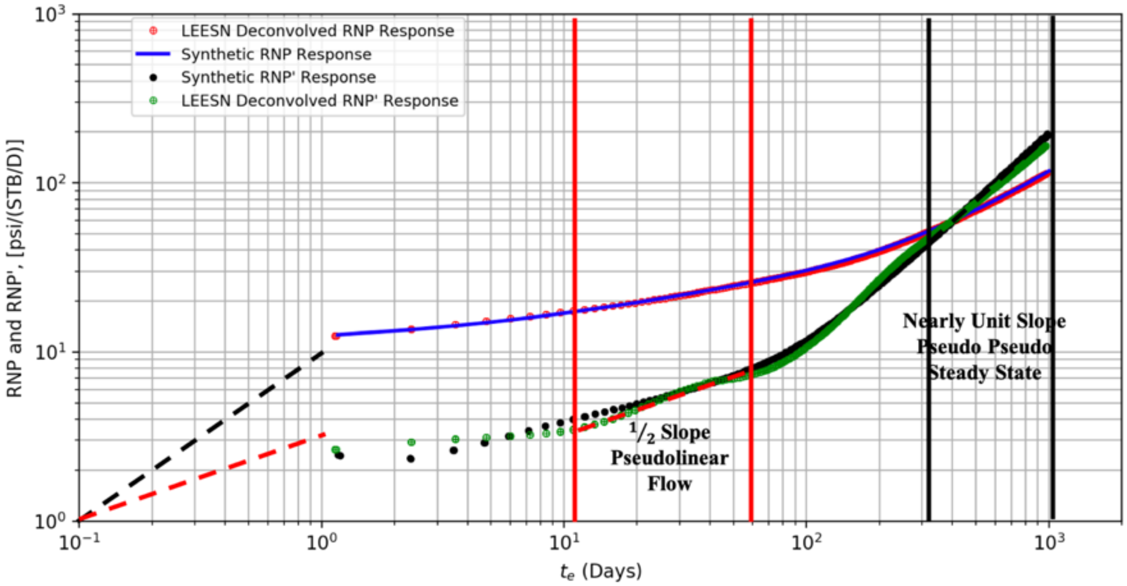


Fig. 53-LEESN deconvolved response for the synthetic noisy-data MTFW case under piecewise BHP control compared to simulated rates on diagnostic RNP, RNP′ vs t_e plot.

The initial production rate in the noisy-data MTFW case is much higher than in the single-fracture case and declines rapidly. We simulated a two-year production period to compare the forecasting performance in both daily production and cumulative production using LEESN. The first-year data is exposed to the model and the forecasted one-year data is shown in red beyond the vertical dashed line in Fig. 54.

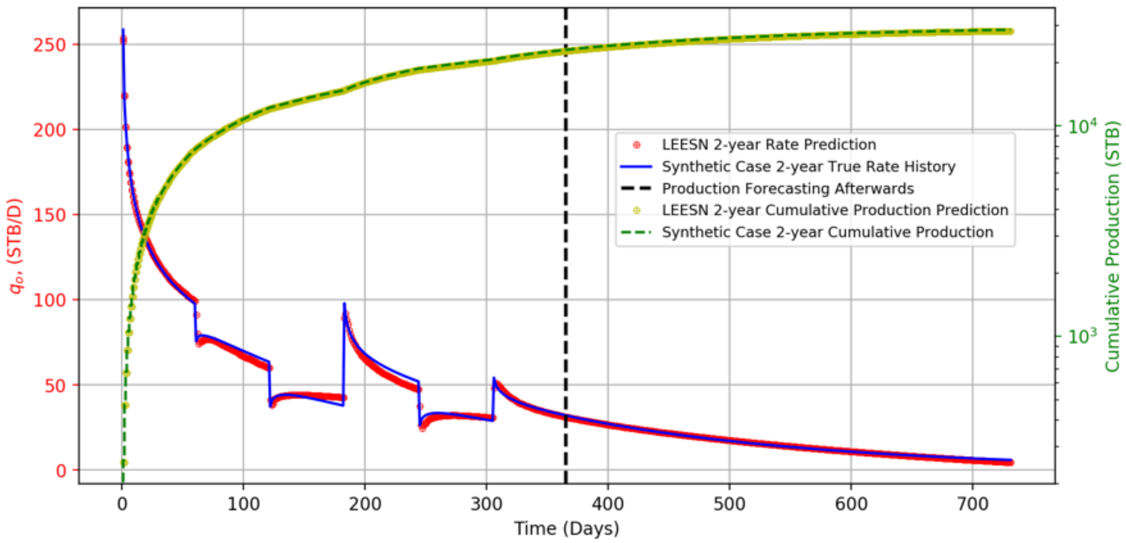


Fig. 54-The comparison of simulated production rate, cumulative production and LEESN forecasted production rate, cumulative production for the noisy-data MTFW case.

3.2 Synthetic Case in Pressure Transient Analysis

3.2.1 Single Fracture Horizontal Well Case

The first synthetic case in pressure transient analysis is for a horizontal well with a single hydraulic fracture in a homogeneous reservoir which we described in section 3.1.1. Unlike the previous setup, we ran the simulation for one year under piecewise surface flow rate (STO) control. The nonlinearity of flow rates with time is guaranteed, so that RNP or PNR functions are not applicable in this scenario. The reservoir flow regimes during the one-year simulation are clearly identified in Fig. 55, and the corresponding simulation parameters are summarized in Table 6.

Part of section 3.2 is reprinted with permission of Society of Petroleum Engineers from Pan, Y., Deng, L., Lee, J. 2019. Data-Driven Deconvolution Using Echo-State Networks Enhances Production Data Analysis in Unconventional Reservoirs. Proceedings of SPE Eastern Regional Meeting, Charleston, West Virginia, 15-17 October. SPE-196598-MS. DOI: 10.2118/196598-MS. Permission conveyed through Copyright Clearance Center, Inc.

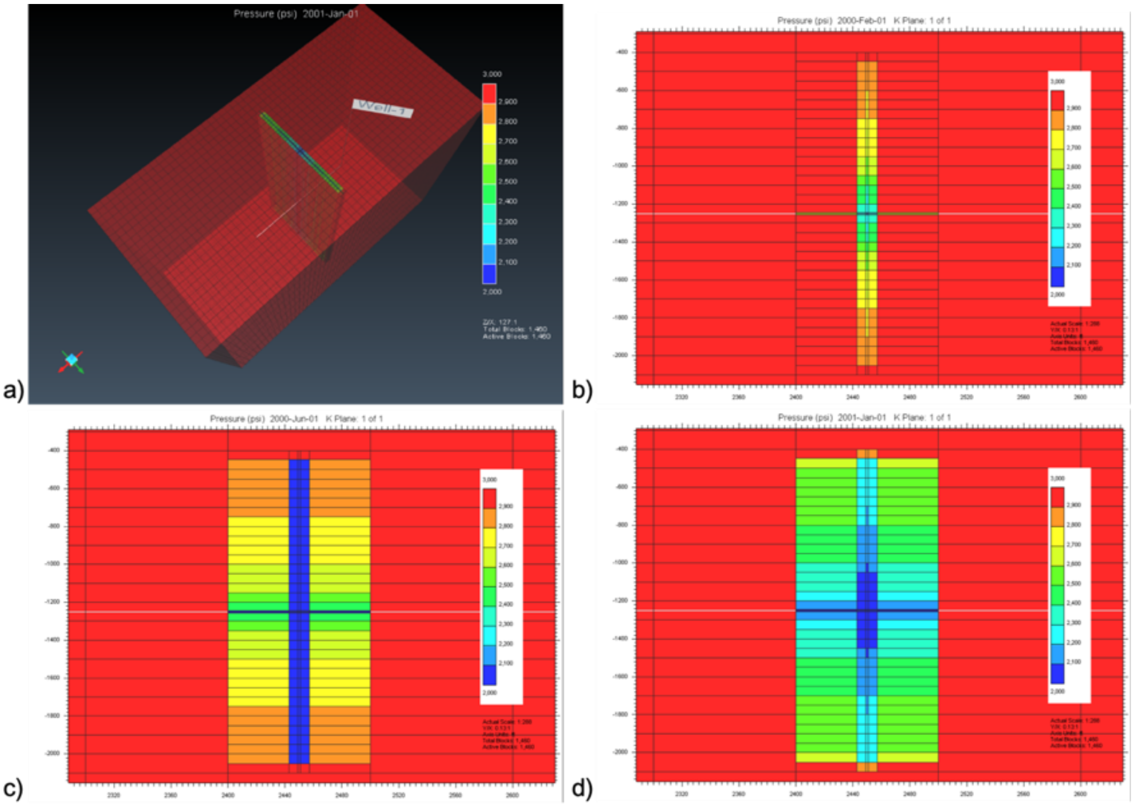
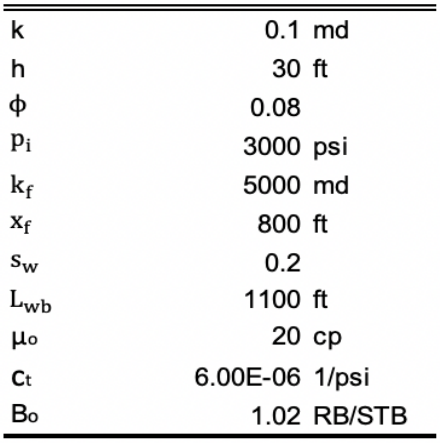


Fig. 55 (a)-3D view of the synthetic single-fracture case pressure map at the end of simulation under piecewise STO; (b)-Synthetic single-fracture case pressure map at the end of 3rd month under piecewise STO; (c)-Synthetic single-fracture case pressure map at the end of the 8th month under piecewise STO; (d)-Synthetic single-fracture case pressure map at the end of 1st year under piecewise STO.



k	0.1	md
h	30	ft
ϕ	0.08	
p_i	3000	psi
k_f	5000	md
x_f	800	ft
s_w	0.2	
L_{wb}	1100	ft
μ_o	20	cp
c_t	6.00E-06	1/psi
B_o	1.02	RB/STB

Table 6-Synthetic Single Fracture Case Simulation Parameters Under Piecewise STO.

Based on the physics training features in Table 3, the critical factor in applying deconvolution in PTA successfully is to reconstruct the training features. Table 7 summarizes the training features for PTA.

Superposition on Rate	$\Sigma q(f) = \sum_{j=1}^{\mathrm{N_t}} [q_j - q_{j-1}]$
Superposition on Linear Time	$\Sigma \Delta t_L(f) = \sum_{j=1}^{\mathrm{N_t}} [q_j - q_{j-1}][\sqrt{t_{\mathrm{N_t}} - t_{j-1}}]$
Superposition on Bilinear Time	$\Sigma \Delta t_{BL}(f) = \sum_{j=1}^{\mathrm{N_t}} [q_j - q_{j-1}]\left[(t_{\mathrm{N_t}} - t_{j-1})^{\frac{1}{4}}\right]$
Superposition on Pseudo Radial Time	$\Sigma \Delta t_{PR}(f) = \sum_{j=1}^{\mathrm{N_t}} [q_j - q_{j-1}][\log(t_{\mathrm{N_t}} - t_{j-1})]$
Superposition on Pseudo Steady-State Time	$\Sigma \Delta t_{PSS}(f) = \sum_{j=1}^{\mathrm{N_t}} [q_j - q_{j-1}][t_{\mathrm{N_t}} - t_{j-1}]$

Table 7-Superposition Time Functions for PTA.

The training features are modified using Eq. 61 and Eq. 59:

$$\mathbf{u_i}(t) = \left\{ \begin{matrix} 1 \\ \Sigma \Delta q \\ \Sigma \Delta t_L \\ \Sigma \Delta t_{BL} \\ \Sigma \Delta t_{PR} \\ \Sigma \Delta t_{PSS} \end{matrix} \right\}, i = 1, \dots, \mathrm{N} \qquad (62)$$

The target $\mathbf{y}(t)$ is simply the drawdown pressure array, a function of t,

$$\mathbf{y}(t) = \{\Delta p(t)\} \qquad (63)$$

Fig. 56 and Fig. 57 shows that the STO was altered every two months corresponding to the variation of pressure drawdown. We added 3% additive white

Gaussian noise to both pressure and rate signals, shown in blue. Both transformed noisy-data signals using Eq. 62 and Eq. 63 were fed into the LEESN algorithm for training purpose, then, after training was completed, the clean rate signal was transformed and input into the proxy model to regenerate a clean pressure drawdown response as shown in Fig. 57 with the red dashed curve. We deconvolved the data based on the well-trained model.

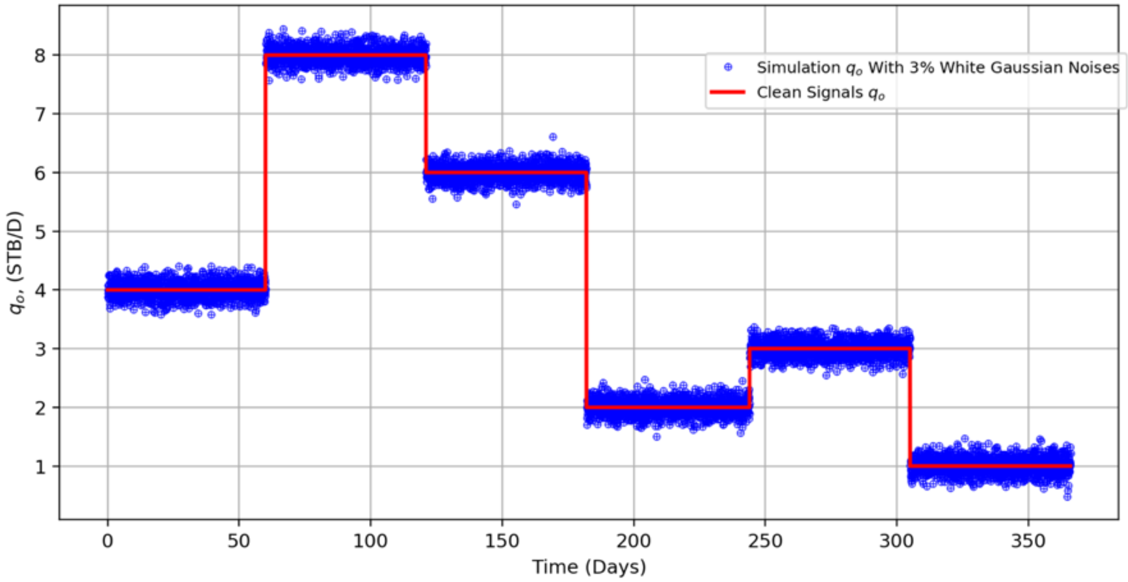


Fig. 56-Clean production rates and noisy rates with 3% additive white Gaussian noise over time for synthetic case I under piecewise STO control.

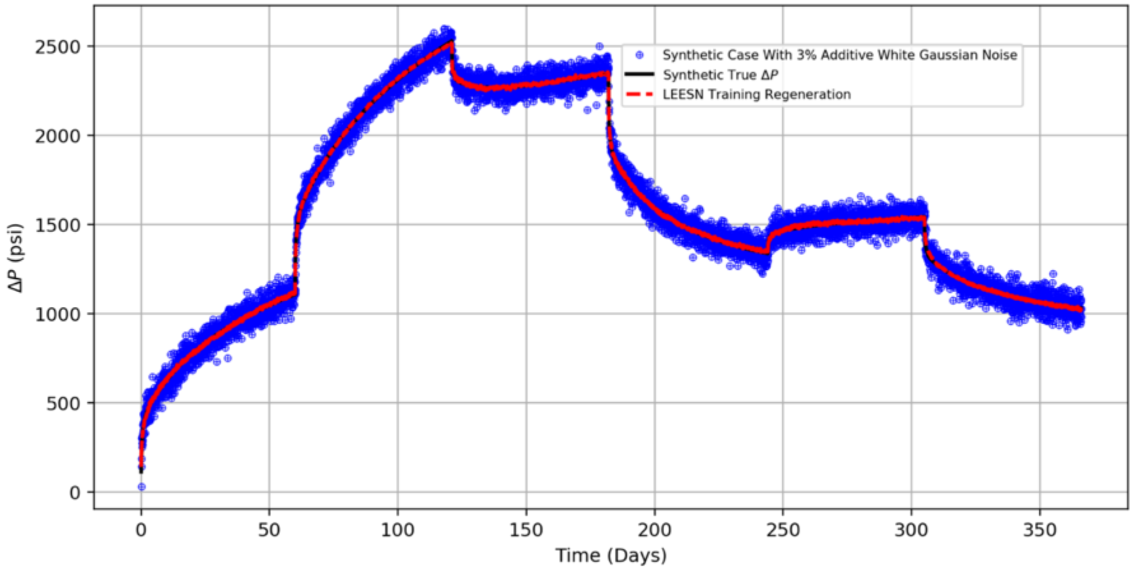


Fig. 57-Simulated pressure drawdown, pressure drawdown with 3% additive white Gaussian noise and model smoothed drawdown pressure regeneration over time for synthetic case I under piecewise STO control.

Deconvolution process is performed by feeding in a constant-rate physics-assisted transformed features using Eq. 62. The resultant pressure drawdown is illustrated in Fig. 58. The deconvolved pressure derivatives are aligned with the simulated pressure derivatives.

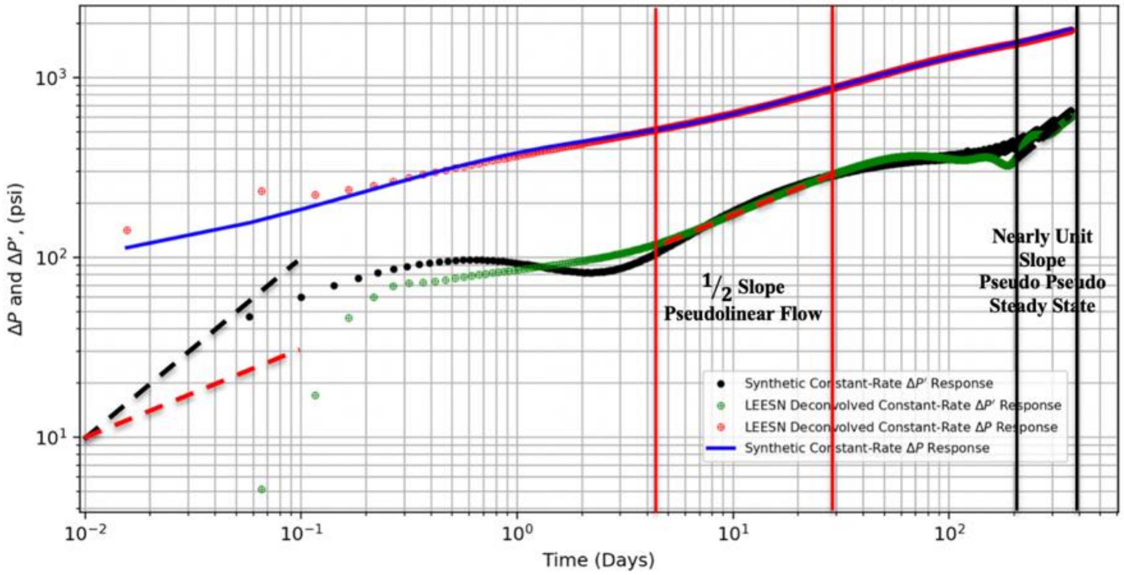


Fig. 58-Synthetic single fracture case production analysis using LEESN deconvolved drawdown pressure response and its Bourdet derivative on a diagnostic plot.

The diagnostic plot presented here conform to the 'basement' reservoir model proposed by Kuchuk, in which the matrix permeability is unlikely to contribute to any hydrocarbon extraction (Kuchuk et al. 2012). In the synthetic single-fracture case, the noisy variable-rate drawdown test from a horizontal well with a single fracture is deconvolved into a constant rate drawdown pressure response. The derivative exhibits a half slope during the pseudo-linear flow period, followed by an ambiguous nearly unit slope possibly corresponding to the pseudo pseudo-steady state period which indicates the 'basement' effects have been initiated. This further demonstrates that physics-based training features and training algorithms play the key roles in identifying the flow regimes even with noisy data.

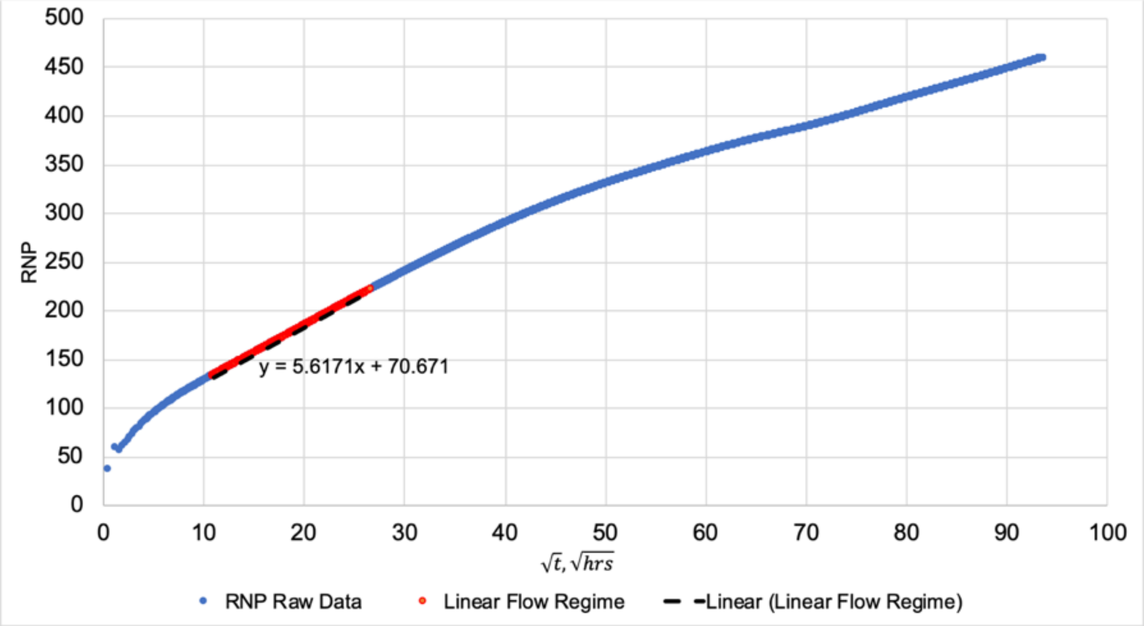
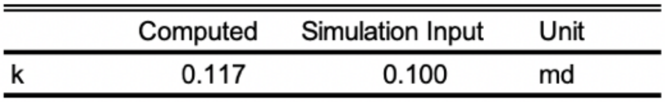


Fig. 59-Synthetic single-fracture case production analysis using LEESN deconvolved drawdown pressure response on a RNP $vs. \sqrt{t}$ specialized plot.

Another purpose of production analysis is to identify flow patterns based on the deconvolved response and determine the corresponding reservoir properties. We use reservoir matrix permeability as an example. In Fig. 59, the linear flow regime is identified; accordingly, Eq. 64 and the reservoir properties in Table 6 allow us to estimate the matrix permeability using the slope m=5.617.

$$x_f\sqrt{k} = 4.064\frac{B\mu}{mh\sqrt{\phi\mu \mathrm{c_t}}} \tag{64}$$

The comparison between calculated matrix permeability and the permeability used as input in synthetic simulation is illustrated in Table 8. The permeability determined from deconvolved pressure drawdown is very close to the synthetic input value.



	Computed	Simulation Input	Unit
k	0.117	0.100	md

Table 8-Comparison of Calculated Matrix Permeability and Simulation Input Matrix Permeability.

3.2.2 Multiple Transverse Fractures Horizontal Well Case

In the second synthetic case to examine PTA, we extended the previous model to a model for a horizontal well with multiple transverse fractures (MTFW). The synthetic data includes a 3-year simulation under piecewise surface flow rate (STO) control, with rate changes every 6-month as shown in Fig. 61. The corresponding pressure drawdown is illustrated in Fig. 62. The reservoir model and critical reservoir parameters are depicted and summarized in Fig. 60 and Table 9.

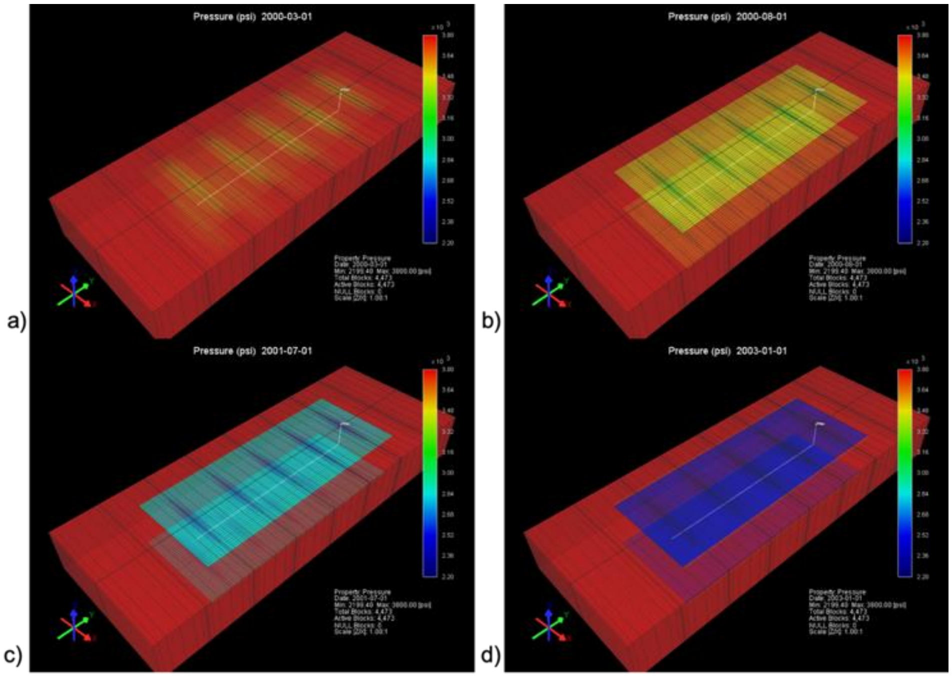


Fig. 60 (a)-Synthetic MTFW case pressure map at the end of 3rd month under piecewise STO; (b)-Synthetic MTFW case pressure map at the end of the 8th month under piecewise STO; (c)-Synthetic MTFW case pressure map at the end of 18th month under piecewise STO; (d)-Synthetic MTFW case pressure map at the end of 36th month under piecewise STO.

k	0.0001	md
k_{SRV}	0.5	md
h	200	ft
ϕ	0.08	
ϕ_{SRV}	0.024	
p_i	3800	psi
k_f	5000	md
x_f	135	ft
n_f	5	
s_w	0.2	
L_{wb}	1000	ft
μ_o	45	cp
c_t	2.00E-06	1/psi
B_o	1.025	RB/STB

Table 9-Synthetic MTFW Case Simulation Parameters Under Piecewise STO.

Again, 3% additive white Gaussian noise was combined with both pressure and rate signals shown in blue. We applied Eq. 62 and Eq. 63 to transform the raw signals into the target training features before we input them into the model. As expected, the alignment between regenerated response and true synthetic pressure drawdown signals can be observed after training was completed.

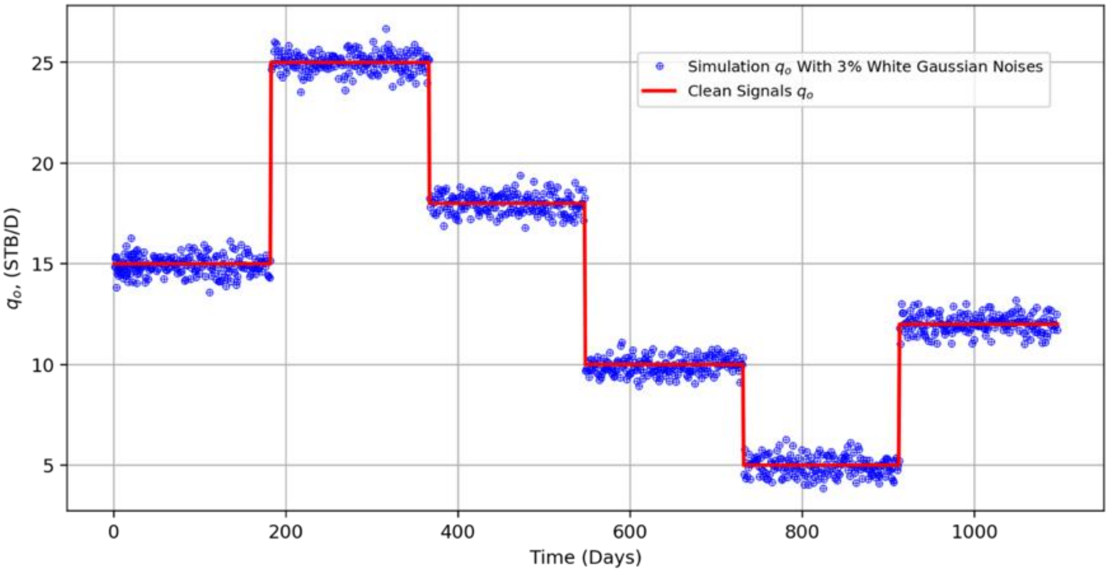


Fig. 61-Clean production rates and rates with 3% additive white Gaussian noise over time for synthetic MTFW case under piecewise STO control.

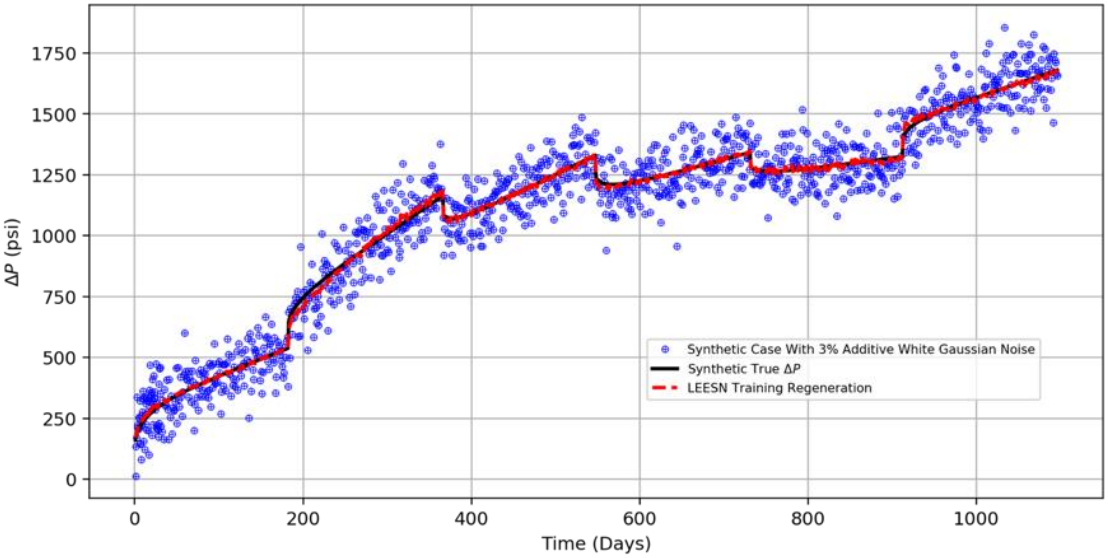


Fig. 62-Simulated pressure drawdown, pressure drawdown with 3% additive white Gaussian noise, and model-smoothed drawdown pressure regeneration over time for synthetic MTFW case under piecewise STO control.

We further input in a set of training features using constant rate to obtain pressure drawdown response. Fig. 63 illustrates the deconvolved results and the corresponding Bourdet derivative. The deconvolved response and the simulated response align well for both pressure drawdown and Bourdet derivative.

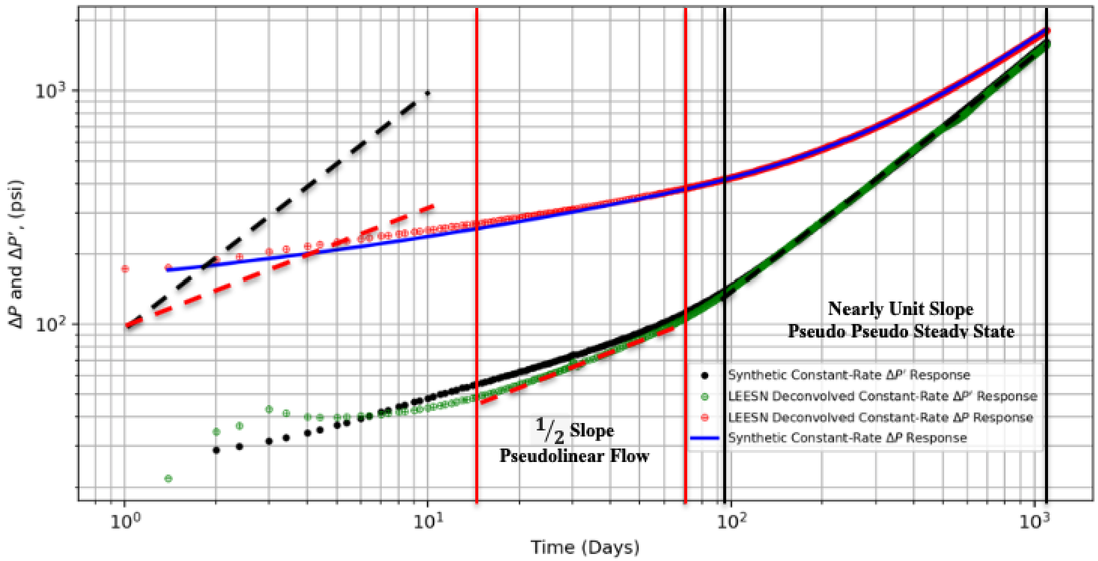


Fig. 63-Synthetic MTFW case production analysis using LEESN deconvolved drawdown pressure response and its Bourdet derivative on a diagnostic plot.

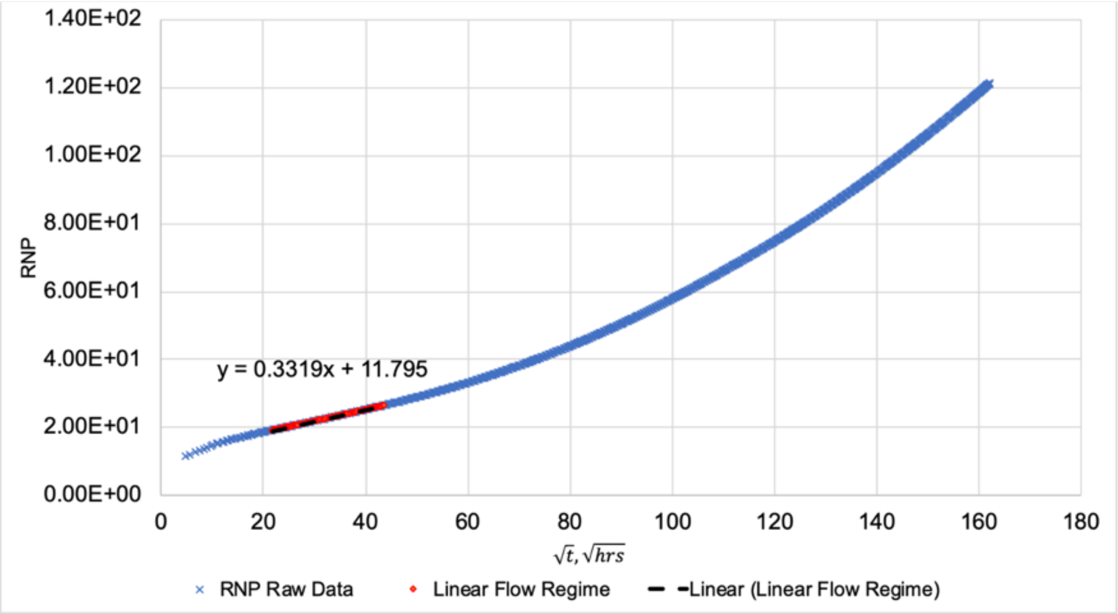


Fig. 64-Synthetic MTFW case production analysis using LEESN deconvolved drawdown pressure response on a RNP $vs. \sqrt{t}$ specialized plot.

The linear flow regime and the pseudo pseudo-steady state flow regime were identified as shown in Fig. 64 and Fig. 65. Reservoir properties were estimated from these plots. Based on Eq. 64, equation for linear flow in a well with multiple fractures has the form of Eq. 65. Eq. 66 is the pseudo pseudo-steady state equation.

$$4n_f x_f \sqrt{k} = 4.064 \frac{B\mu}{mh\sqrt{\phi\mu c_t}} \tag{65}$$

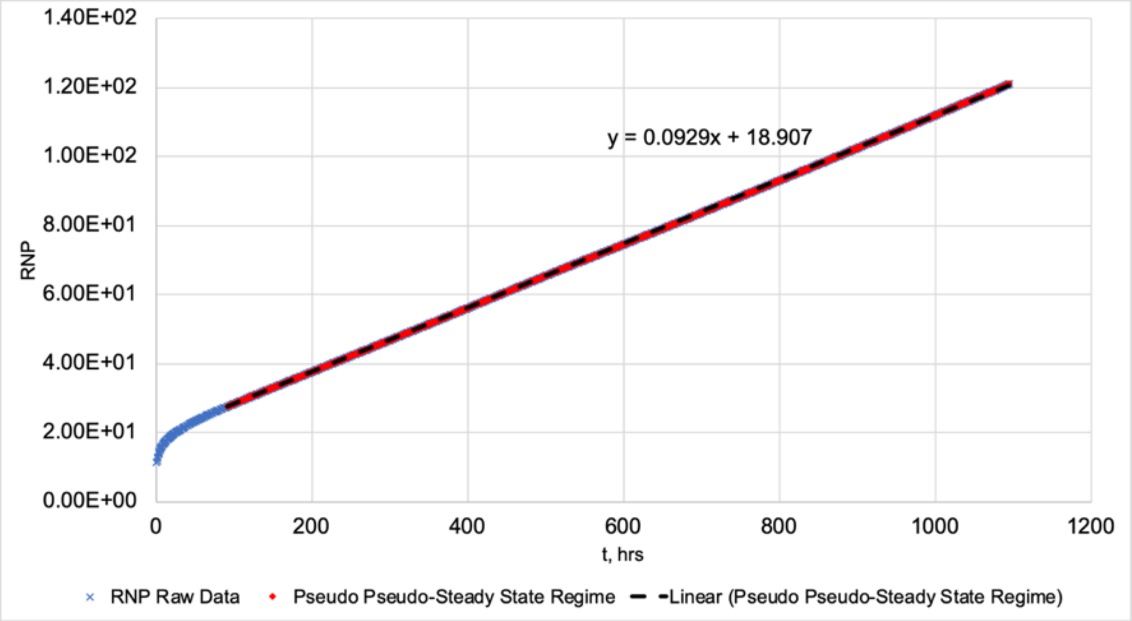
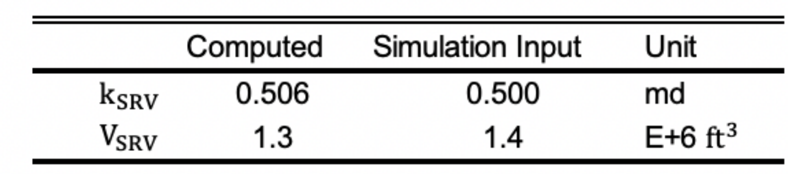


Fig. 65-Synthetic MTFW case production analysis using LEESN deconvolved drawdown pressure response on a RNP $vs. \Delta t$ specialized plot.

$$V_{SRV} = A_{SRV} h\phi = \frac{0.234}{mc_t} \tag{66}$$

Matrix permeabilities and stimulated reservoir volumes are compared in Table 10. The difference between the results are acceptable, further demonstrating that LEESN can perform deconvolution for horizontal wells with single and multiple fractures.



	Computed	Simulation Input	Unit
k_{SRV}	0.506	0.500	md
V_{SRV}	1.3	1.4	E+6 ft^3

Table 10-Comparison of Calculated k_{SRV}, V_{SRV} and Simulation Input k_{SRV}, V_{SRV}.

3.3 Synthetic Case in Production Surveillance

3.3.1 Multiple Transverse Fractures Horizontal Well Case Model Setup

Performance of deep-learning based algorithms in training and validation ordinarily rely on the richness of data. The frequency of data acquisition in a PDG depends on the initial setup and the resolution of the tool. We developed a synthetic case to approximate a PDG recording downhole data hourly for 3.5 years with variable pressures and variable rates throughout the entire sampling period. The synthetic case shown in Fig. 66 and Fig. 67 is for a homogeneous reservoir produced through a horizontal well with multiple transverse fractures.

Part of section 3.3 is reprinted with permission of URTeC from Pan, Y., Bi, R., Zhou, P., Deng L., Lee, J. 2019. An Effective Physics-Based Deep Learning Model for Enhancing Production Surveillance and Analysis in Unconventional Reservoirs. Proceedings of Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July. URTEC-2019-145. DOI:10.15530/urtec-2019-145. Copyright 2019 URTeC.

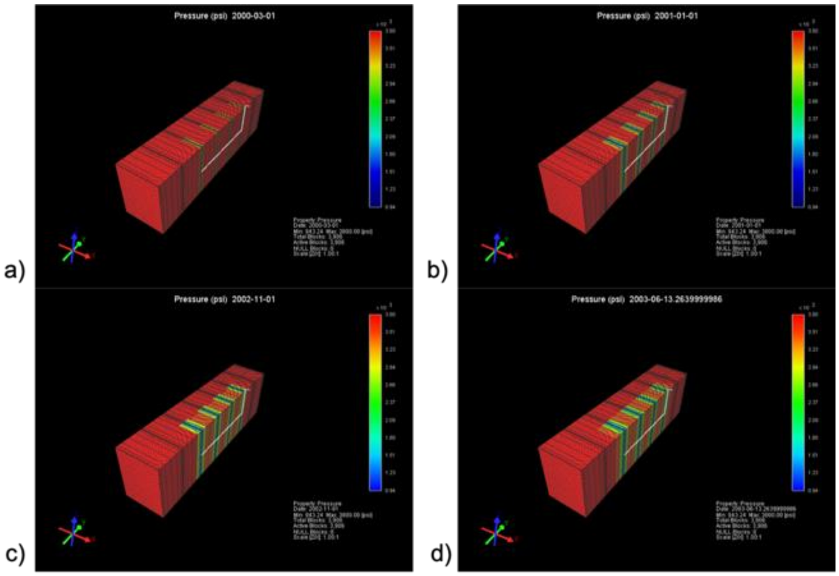


Fig. 66 (a)-Synthetic case reservoir pressure after two months of production; (b)-Synthetic case reservoir pressure at the end of 1st year; (c)-Synthetic case reservoir pressure during fracture interference; (d)-Synthetic case reservoir pressure profile at almost 3.5 years.

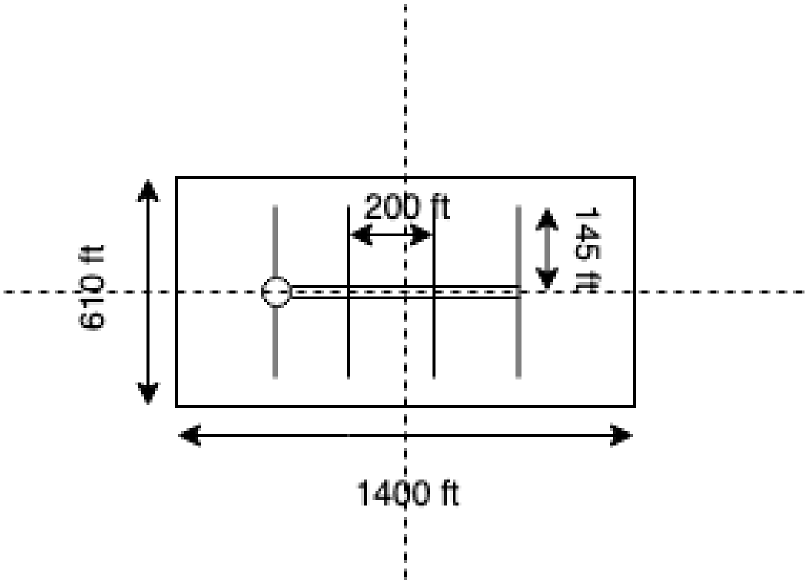
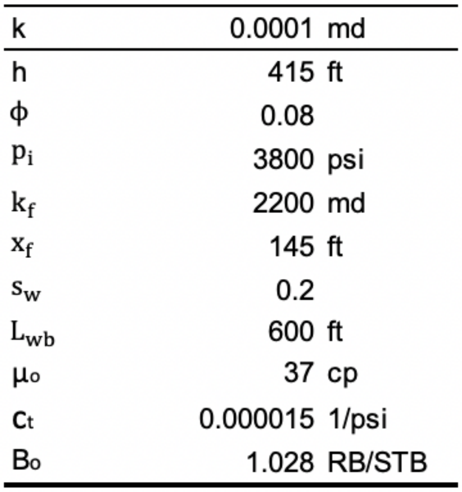


Fig. 67-Top-down view of the reservoir.

Reservoir parameters are summarized in Table 11.



k	0.0001	md
h	415	ft
ϕ	0.08	
p_i	3800	psi
k_f	2200	md
x_f	145	ft
s_w	0.2	
L_{wb}	600	ft
μ_o	37	cp
c_t	0.000015	1/psi
B_o	1.028	RB/STB

Table 11-Synthetic MTFW Case Simulation Parameters Under Piecewise STO.

3.3.2 Scenario I Production Surveillance with Only Production Rates

Scenario I approximates production surveillance data using PDG to record only production rates. In this test, 5% additive white Gaussian noise (WGN), aberrant segments and outliers were artificially added to mimic real data gathered downhole. The noisy rate profile and synthetic rates are shown in Fig. 68 in the blue solid line and green solid line. Additionally, three segments were distorted to form the aberrant segments, and we generated outliers randomly throughout the data sequence. Besides, by removing a 500-data-point sequence shown in red dashed line is to simulate a workover or system offline for approximately 500 hours.

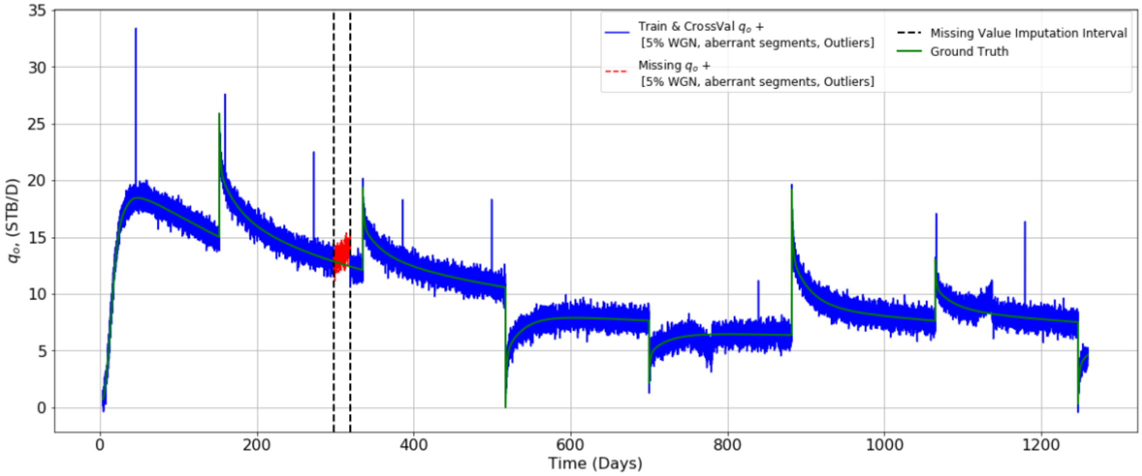


Fig. 68-Synthetic case scenario I – production rates + [5% WGN, aberrant segments, outliers] and 500 consecutive hours of missing values.

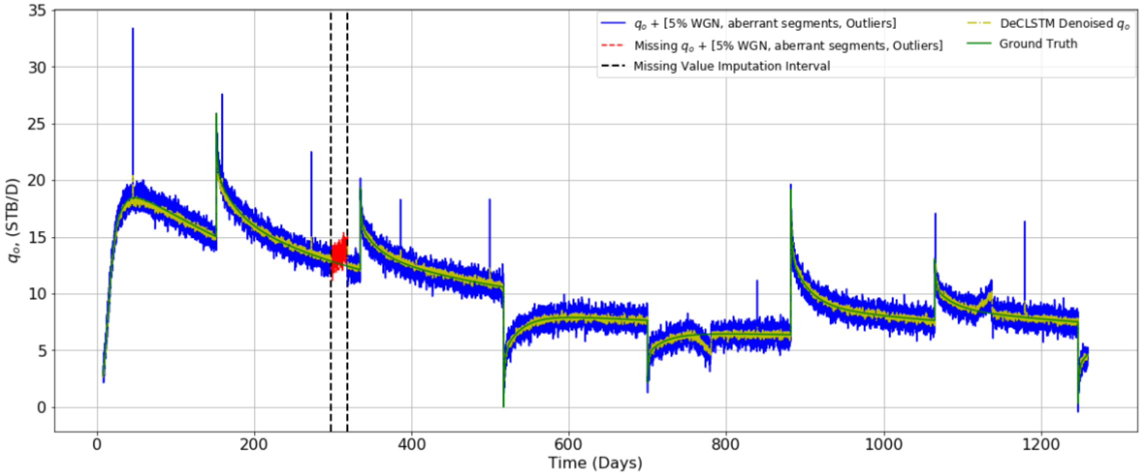


Fig. 69-DeLSTM denoise process for scenario I.

The noisy distorted rate signal was divided into batches of sequences for training and validation. A k-fold cross-validation was then performed and the model was saved. The noisy distorted signal was first reconstructed with the denoising process shown in yellow in Fig. 69. We observe that the noise and outliers were alleviated, but the aberrant segments still prevail in the data sequence. Therefore, the second step was to reconstruct the missing intervals and distorted segments. To simplify the process, we combined the missing interval and one aberrant segment into a special data sequence to be reconstructed as illustrated in red; the duration of the data sequence remains approximately 500 hours.

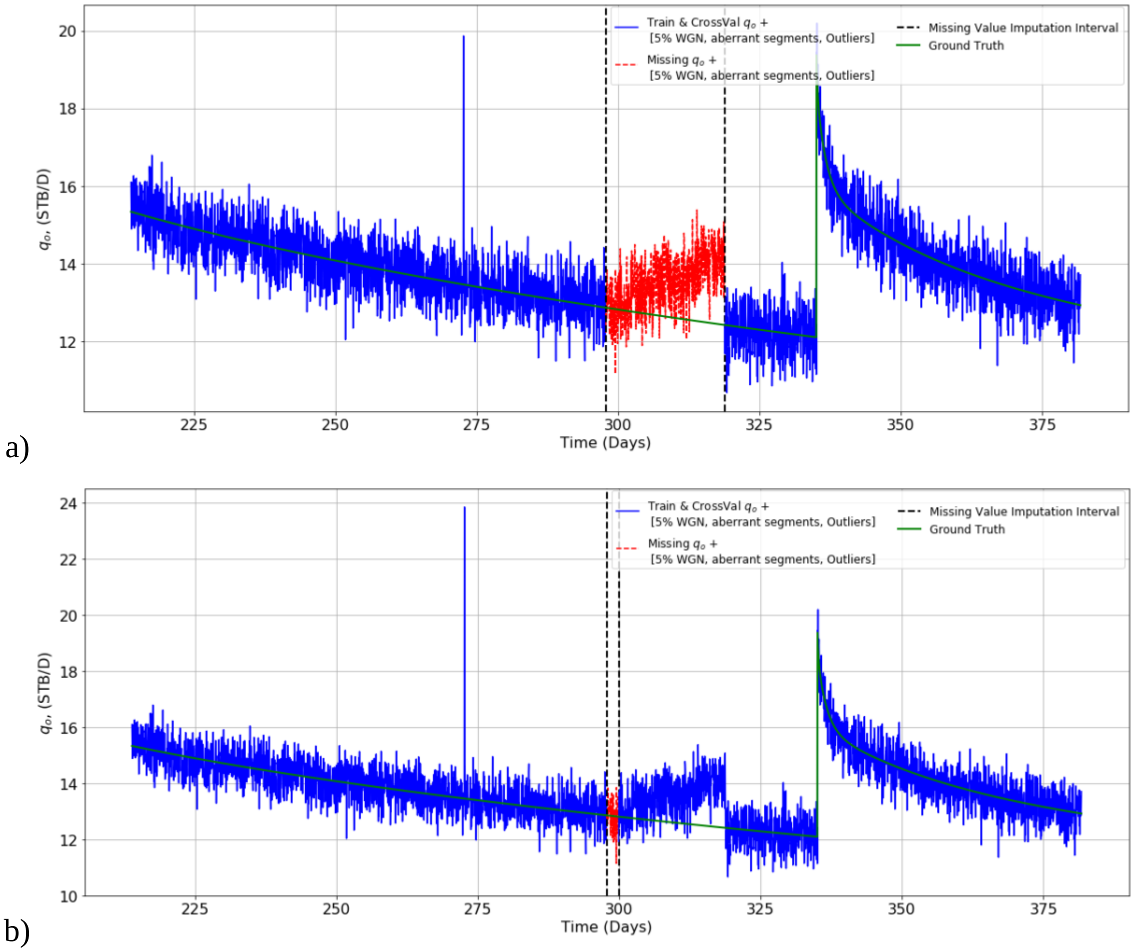


Fig. 70 (a)-Zoom-in view of 500-hour continuous missing data sequence. (b)-First 100 hours missing data sequence from subdivided 500-hour missing interval.

Fig. 70 is a zoom-in view of the missing interval. The missing 500-hour duration data interval was segregated into 5 shorter data sequences. The window length is then 100 data points. By using the nature of the rolling window and flexibility of the DeLSTM structure displayed in Fig. 17, the first 100 continuous missing values were imputed, and the result is displayed in Fig. 71. The ground-truth synthetic rates and the imputed values align well.

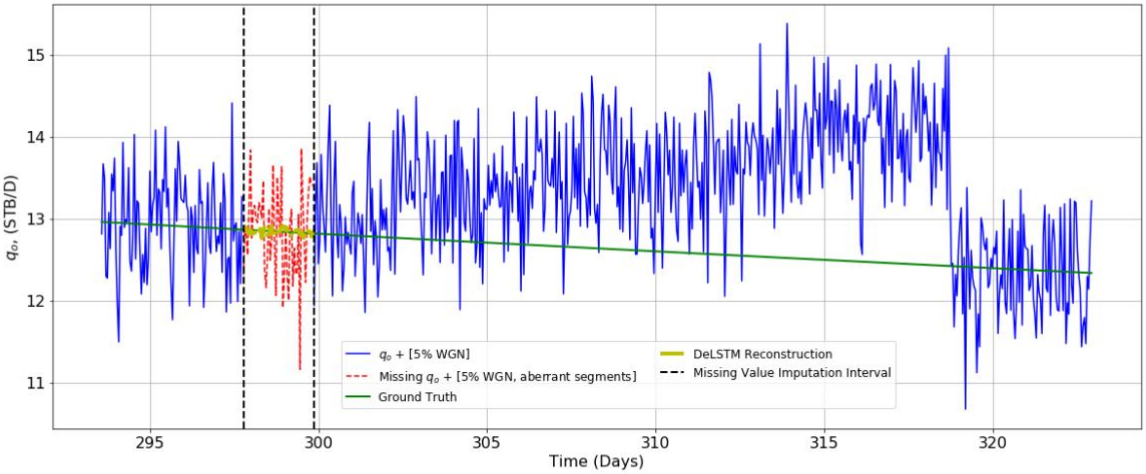


Fig. 71-First imputed rates for first 100 hours of missing value compared to ground-truth synthetic rates and noisy aberrant rates.

We then reconstructed the remaining 400-data-point sequence, and the imputed 500-hour duration missing interval shown in Fig. 72. The deviation from the ground-truth production rates is enlarged as the rolling window moves forward. This result was expected, since the prediction for the first 100 hours is based on the previously observed sequence, and as the rolling window moves more deeply, the later data sequence depends solely on the previously predicted values and the error increases. More importantly, since only production rates are recorded during production surveillance, DeLSTM is used only to discover the time dependence between data sequences without considering any physics. This indicates the limitation of using DeLSTM for long-term production forecasting.

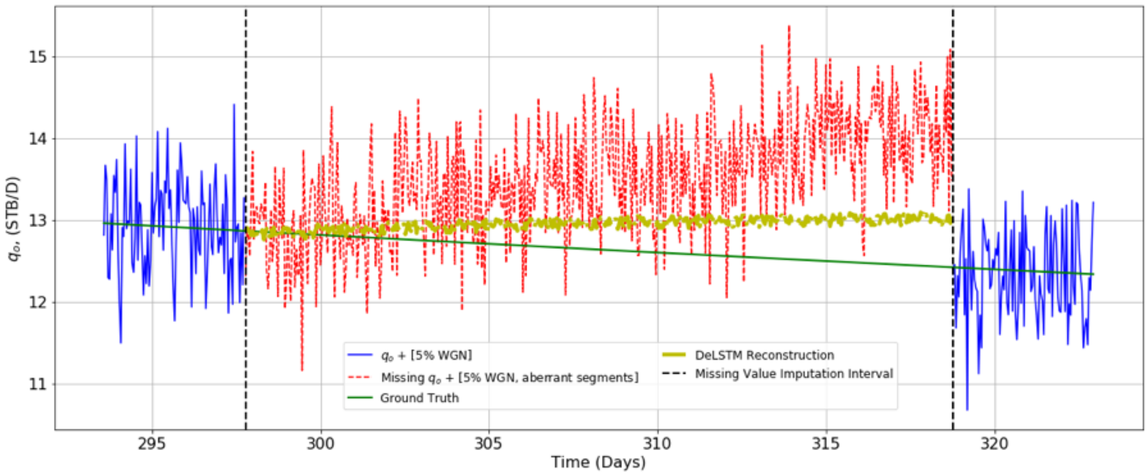


Fig. 72-DeLSTM imputation for 500-hour duration continuous data compared with ground truth production rates and noisy aberrant rates.

3.3.3 Scenario II Production Surveillance with Both BHP and Production Rates

Scenario II mimics production monitoring based on surveying both bottom hole pressures (BHP) and production rates. The ground-truth synthetic production rates are simulated by specifying piecewise BHP which changes every 6 months as shown in Fig. 73. An approximately 5000-hour missing interval was generated by removing the data sequence shown by the red dashed curve. Both signals outside of the missing interval are exposed to the model for training using Eq. 58 as training inputs and Eq. 59 as training outputs. The purpose of this clean signal test is to demonstrate the applicability of these physics-based training features.

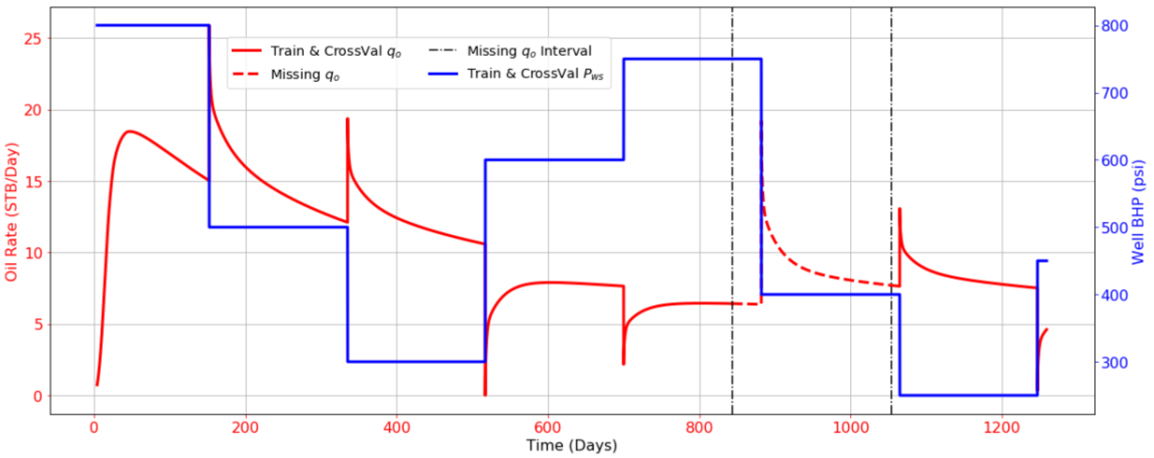


Fig. 73-Synthetic case scenario II – clean production rates with 5000-hour of consecutive missing values and corresponding clean BHPs.

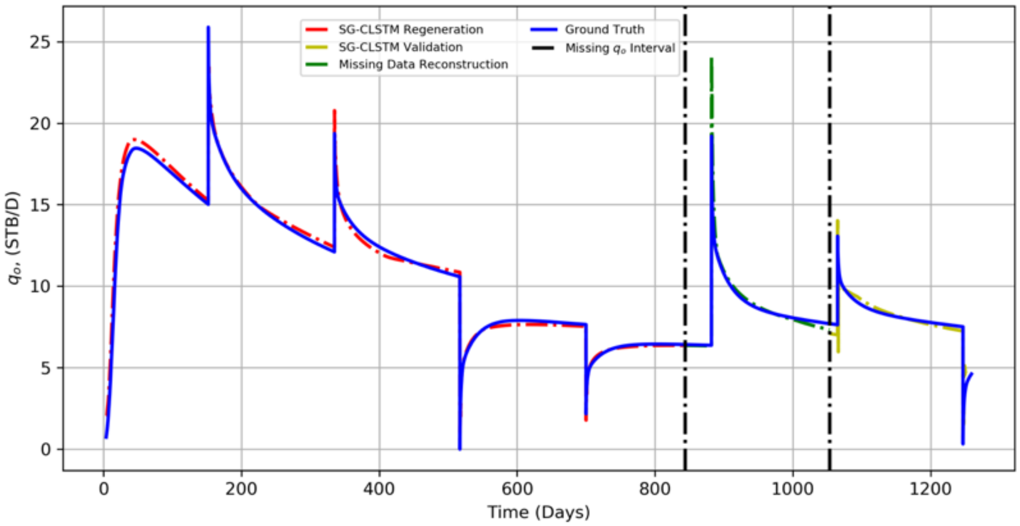


Fig. 74-Comparison of reconstructed production rates and ground-truth synthetic production rates.

The dashed curves in Fig. 74 indicates the reconstructed production profile. The blue solid line is the target output. This figure indicates that the green dashed curve representing the missing interval regeneration is in acceptable alignment with the ground-truth signal and that the matches in training and validation are good. This demonstrates the critical importance of using physics-based training features. We further tested these features by specifying a constant pressure drawdown to deconvolve the variable-pressures variable-rate data into a constant pressure rate response, shown in Fig. 75. The good alignment of constant-pressure rate response and good matches of cumulative production further demonstrate the versatility of our proposed algorithm.

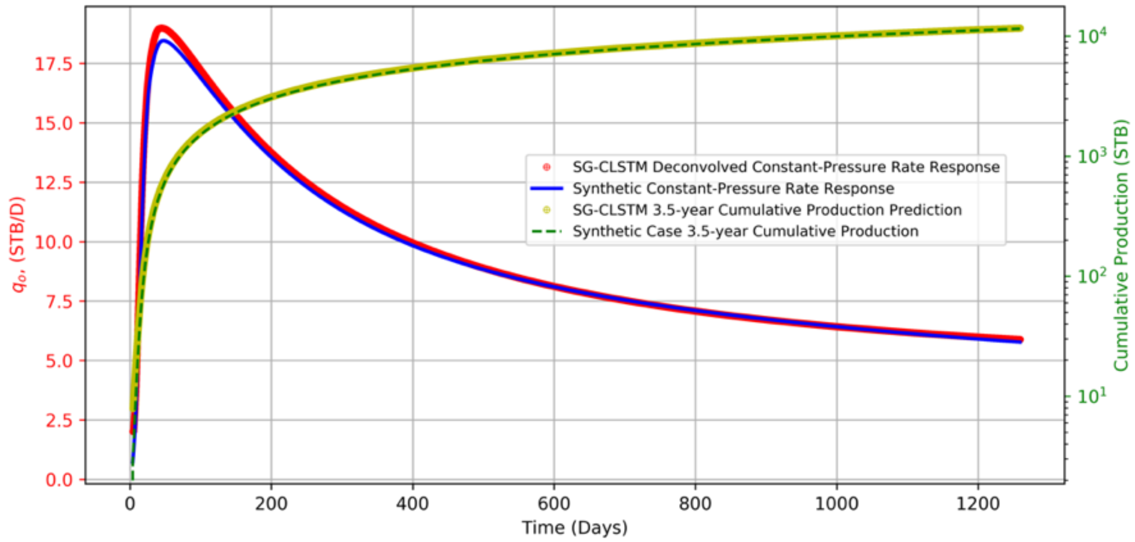


Fig. 75-Comparison of simulated rates for constant pressure drawdown and corresponding cumulative production with SG-CLSTM deconvolved production rates and cumulative production for clean base case.

As in scenario I, we added 5% additive WGN, aberrant segments and outliers to the clean signals to construct a noise-distorted case, illustrated in Fig. 76 and Fig. 77. The same 5000-hour continual missing interval is hidden for model imputation.

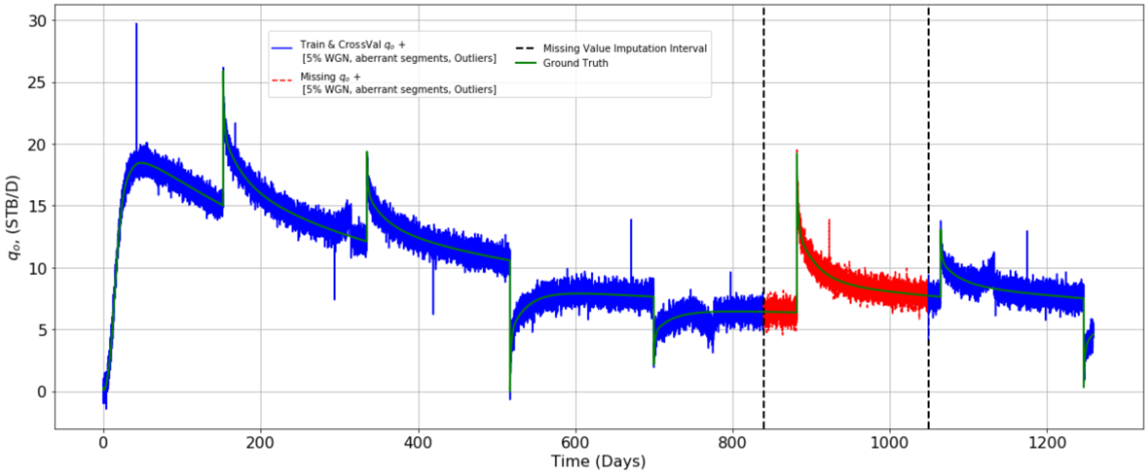


Fig. 76-Synthetic case scenario II – production rates + [5% WGN, aberrant segments, outliers] with 5000 consecutive hours of missing values.

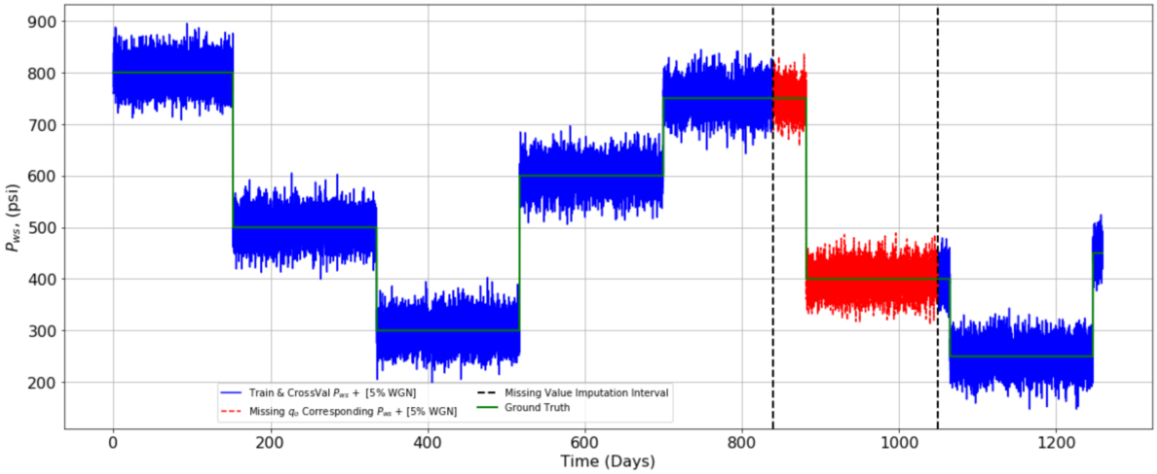


Fig. 77-Synthetic case scenario II – Corresponding BHP + [5% WGN].

By following the systematic workflow illustrated in Fig. 18, both signals were filtered by a SG polynomial smoothing algorithm; results are displayed in Fig. 78 and Fig. 79. Obviously, the SG algorithm can smooth the noisy signals with outliers, however, the aberrant segments remained problematic.

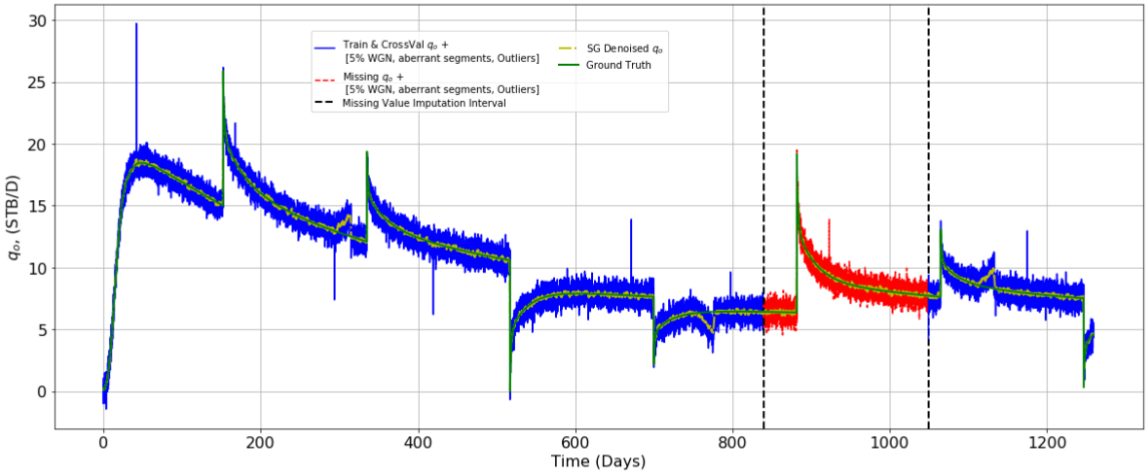


Fig. 78-Synthetic case scenario II – SG smoothed production rate signals compared to ground-truth synthetic rates + [5% WGN, aberrant segments, outliers].

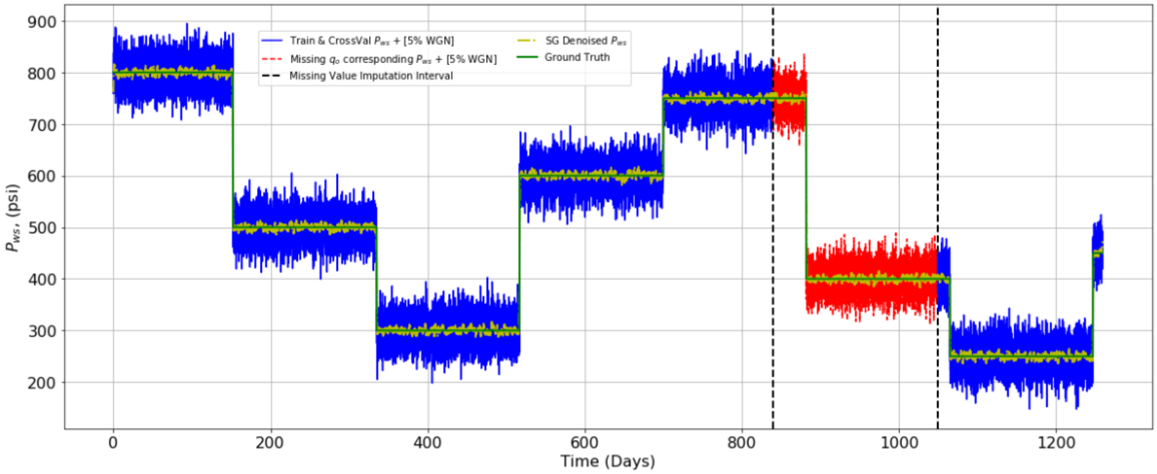


Fig. 79-Synthetic case scenario II – corresponding SG smoothed pressure signals compared to ground-truth synthetic pressures + [5% WGN].

Both denoised signals were sent to the CLSTM for training and cross-validation. The rate signals were reproduced and the missing interval was reconstructed as presented in Fig. 80. We can observe that SG-CLSTM is able to reconstruct the missing flow rates and simultaneously alleviate the distorted signals.

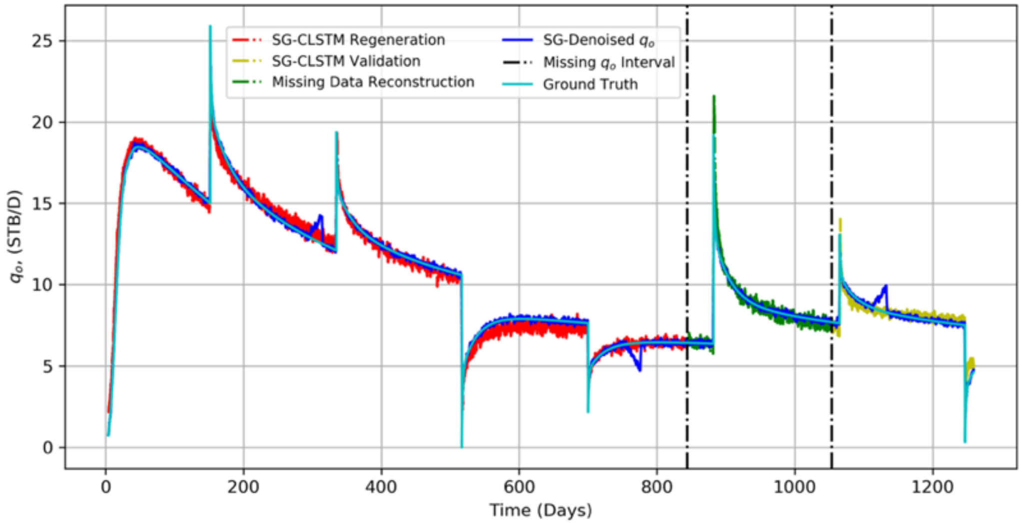


Fig. 80-Comparison of reconstructed, SG-denoised and ground-truth synthetic production rates for 5% noisy distorted case.

Again, a constant-pressure-drop was transformed using Eq. 58 and fed in to obtain the rate response. Fig. 81 illustrates the deconvolved daily production rates and cumulative production. We see some deviation from the synthetic true signals, however, the overall prediction is acceptable. Thus, we have demonstrated that SG-CLSTM is able to deconvolve the variable-pressure, variable-rate data into a constant-pressure-drawdown rate response for further production analysis and forecasting and to reconstruct noisy signals with aberrant segments, outliers, and missing values.

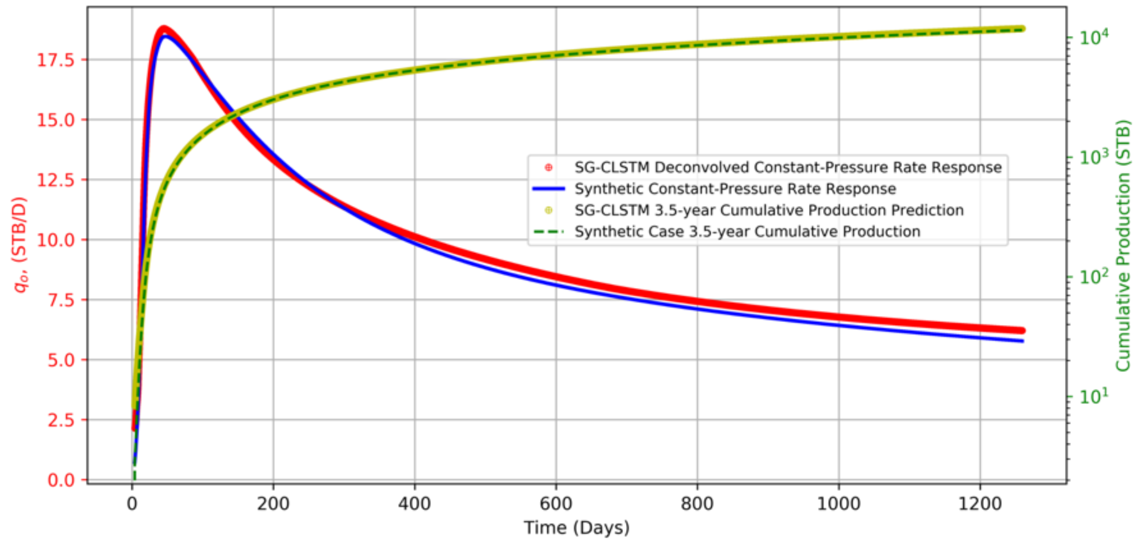


Fig. 81-Comparison of constant pressure drawdown from simulated production rates and simulated cumulative production to SG-CLSTM deconvolved production rates and cumulative production for 5% noise-distorted case.

CHAPTER IV

APPLICATION TO FIELD CASE

4.1 Forecasting Performance Using LEESN

After applications in the synthetic cases described in section 3.1 were completed, we analyzed a field case with about 1-year of production history, including rates and bottom hole pressures. Since no data or supplementary information was available to validate deconvolution, we utilized LEESN only for hindcasting to assess its forecasting capability.

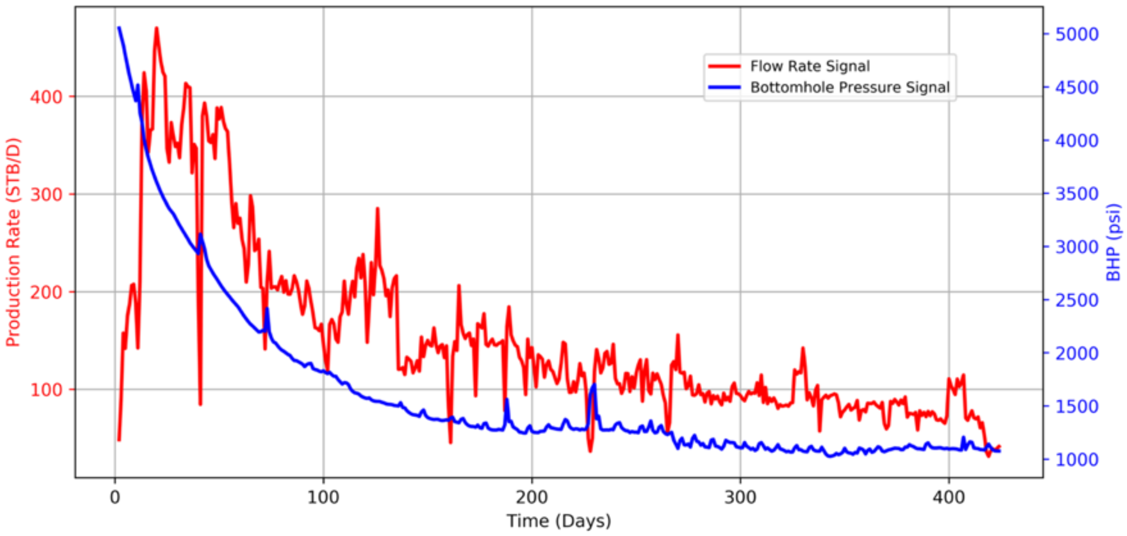


Fig. 82-Bottom hole pressure and flow rate histories for field case 50H.

Part of Chapter IV is reprinted with permission of Society of Petroleum Engineers from Pan, Y., Zhou, P., Deng, L., Lee, J. 2019. Production Analysis and Forecasting for Unconventional Reservoirs Using Laplacian Echo-State Networks. Proceedings of SPE Western Regional Meeting, San Jose, California, 23-26 April. SPE-195243-MS. DOI:10.2118/195243-MS, and from Pan, Y., Deng, L., Lee, J. 2019. Data-Driven Deconvolution Using Echo-State Networks Enhances Production Data Analysis in Unconventional Reservoirs. Proceedings of SPE Eastern Regional Meeting, Charleston, West Virginia, 15-17 October. SPE-196598-MS. DOI: 10.2118/196598-MS. Permission conveyed through Copyright Clearance Center, Inc.; and permission of URTeC from Pan, Y., Bi, R., Zhou, P., Deng L., Lee, J. 2019. An Effective Physics-Based Deep Learning Model for Enhancing Production Surveillance and Analysis in Unconventional Reservoirs. Proceedings of Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July. URTEC-2019-145. DOI:10.15530/urtec-2019-145. Copyright 2019 URTeC.

Rates and pressure changes from test well is 50H (in an unconventional reservoir) shown in Fig. 82. Rate changes are uncorrelated with pressure changes. The history is divided into 20% training and 80% forecasting segments.

Fig. 83 shows the forecasting results. We observe a good alignment in overall tendencies which demonstrates the effectiveness of our proposed methodology in the field case. The physics-based training features and training algorithm, honoring the transient low mechanism, has been further validated. We then predicted well performance for an additional 10 years as shown in Fig. 84. The nature of the constant-pressure rate decline provides alternative forecasting method to estimate unconventional reserves. We recommend further investigation based on the model limitations in the near future.

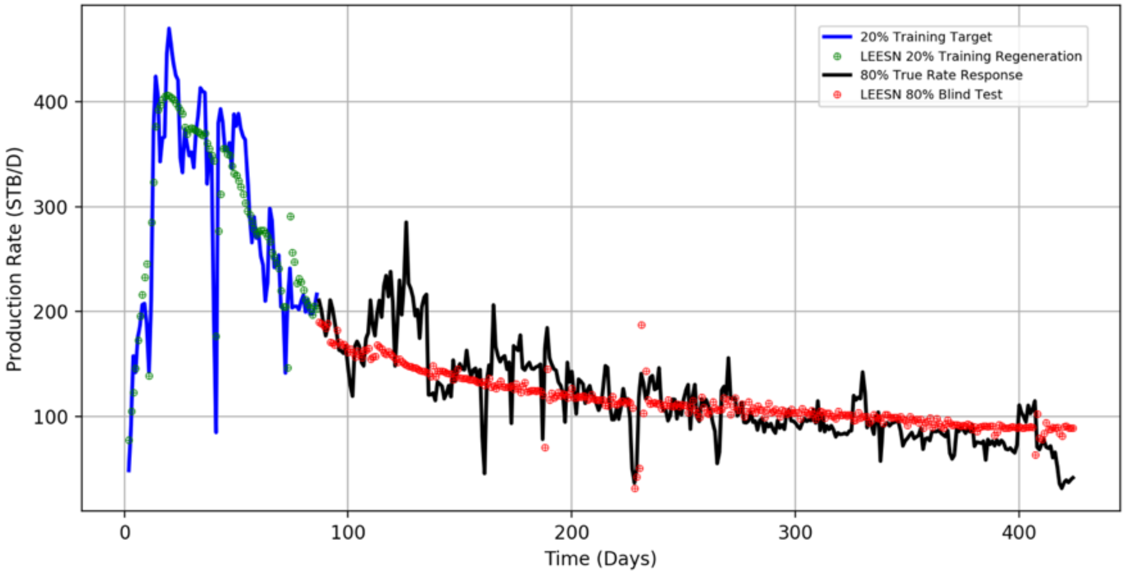


Fig. 83-LEESN hindcast test for well 50H using 20% training, 80% forecasting and comparison with observed flow rates.

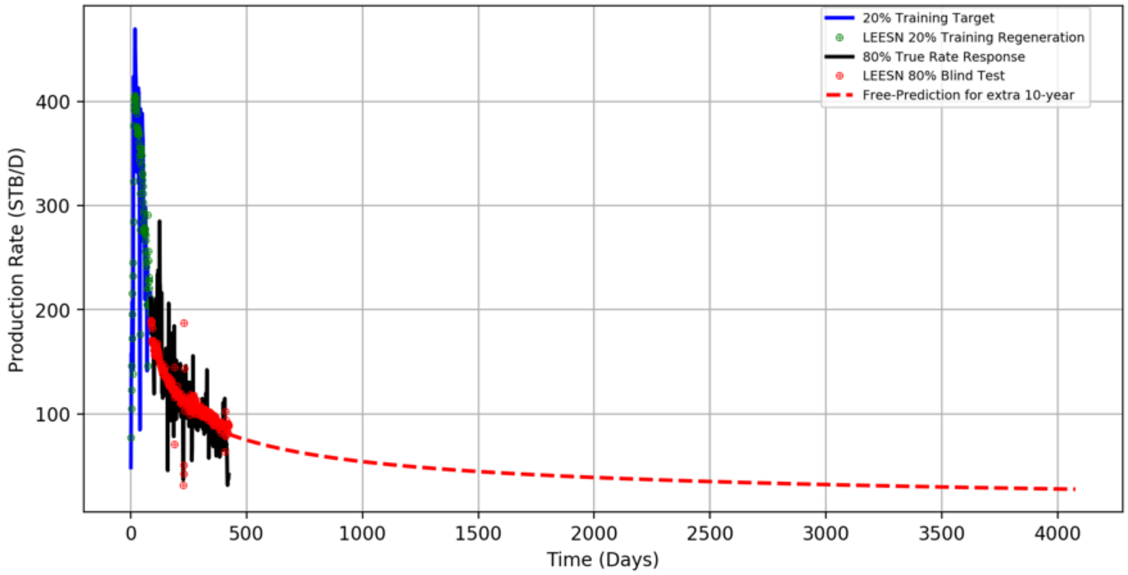


Fig. 84-LEESN hindcast test for well 50H with 10-year rate forecast at constant-BHP.

4.2 Forecasting Performance Using BESN

Proper application of Bayes' theorem requires that we know the probability of the errors to forecast future production correctly. To study this application, we used well 50H to test the versatility of the Bayesian Echo State Network (BESN). As in section 4.1, we divided the data set into 20% training and 80% testing segments, and we further input the same constant pressure drawdown to forecast an extra 10 years production. The hindcast result is shown in Fig. 85 and the forecasting result is illustrated in Fig. 86 with 10% error interval. These results reaffirm that using BESN could provide more confidence in production forecasting.

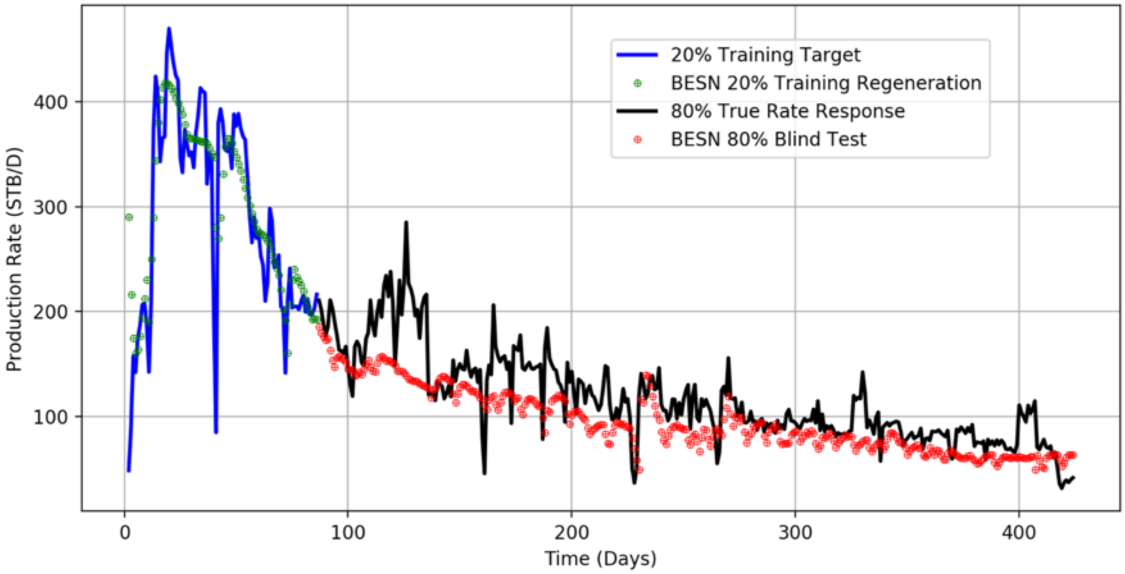


Fig. 85-BESN hindcast test for well 50H using 20% training, 80% forecasting and comparison with observed rates.

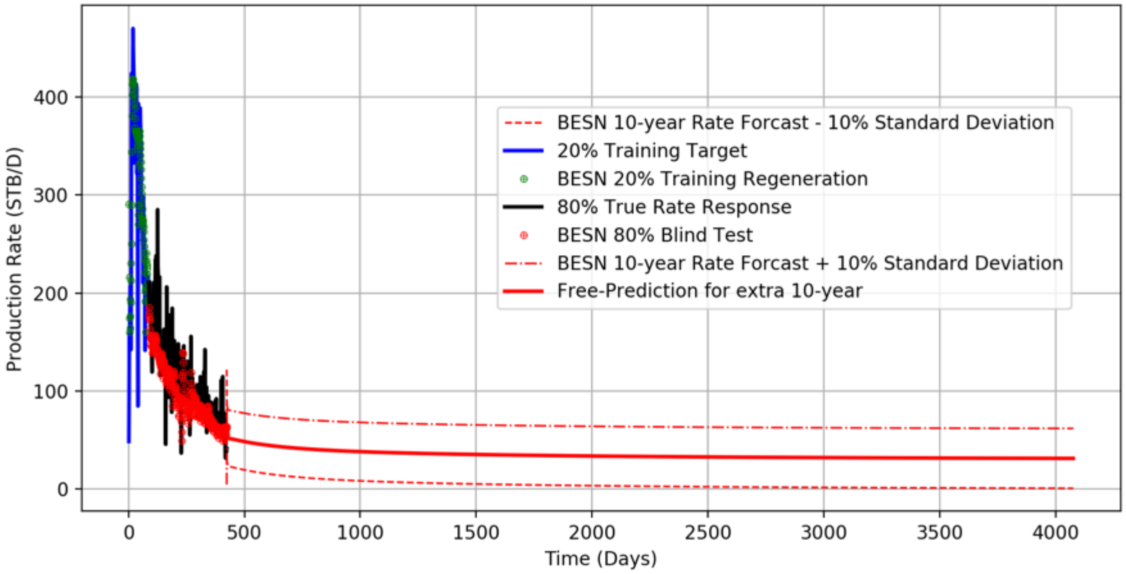


Fig. 86-BESN hindcast test for well 50H and additional 10-year forecast for constant pressure drop with a 10% error interval.

4.3 Deconvolution Performance in PTA Using LEESN

After application to synthetic cases in PTA using LEESN (presented in section 3.2), we studied a field case with nearly 400 consecutive days of pressure and rate records. We applied LEESN for the pressure-rate deconvolution. The raw signals are from a well located in a West Texas unconventional resource play, and both signals are shown in Fig. 87. The highly variable rates and pressure are somewhat uncorrelated indicating noise within the raw data. We used hindcasting to validate the deconvolution performance. The production history was separated into an 80% training segments and a 20% validation segments as illustrated in Fig. 88. The training regeneration is aligned with the observed pressure drawdown.

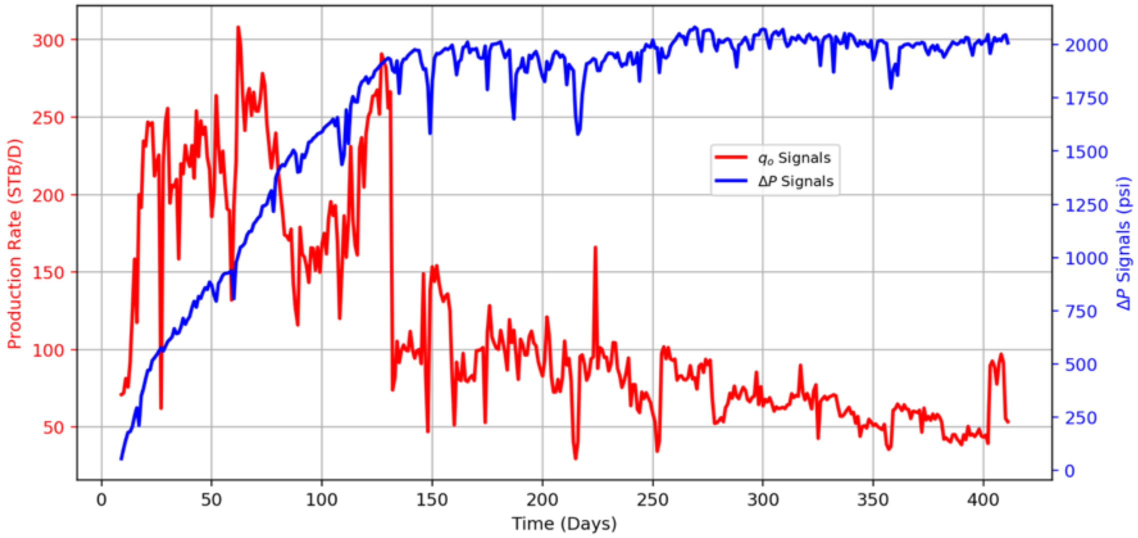


Fig. 87-Pressure drawdown and flow rate histories for well 46H.

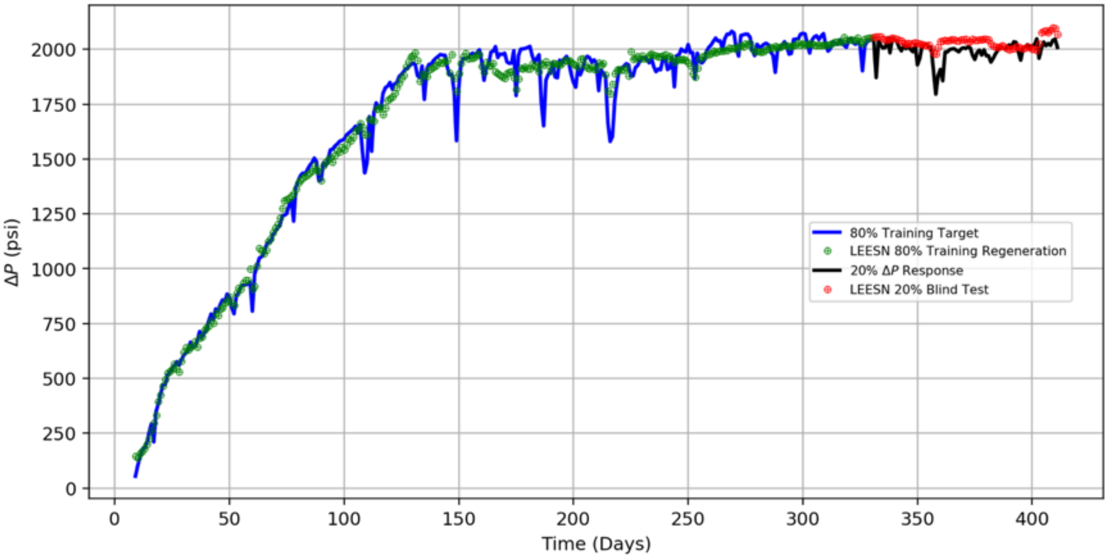


Fig. 88-LEESN hindcast test for well 46H using 80% training, 20% forecasting segments, with comparison to observed pressure drawdown histories.

Fig. 89 shows the deconvolved response using diagnostic plots. We can observe the half-slope pseudolinear flow regime. The following pseudo pseudo-steady state flow regime may be somewhat controversial as the nearly unit slope has not yet prevailed for a long time. Since we do not have any additional information about this well, we do not present an estimation of reservoir properties.

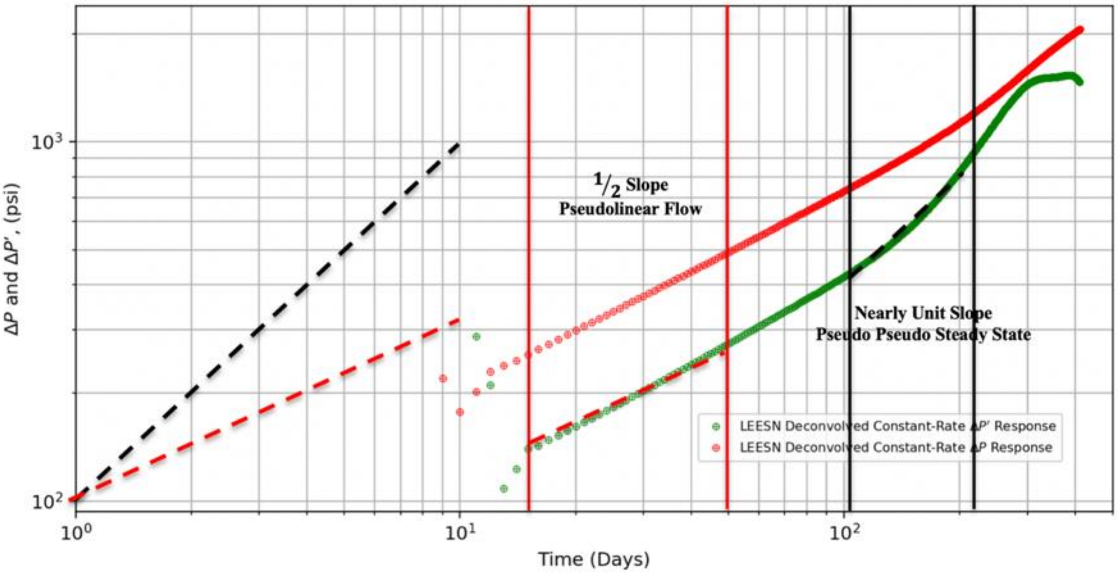


Fig. 89-Well 46H production analysis using LEESN deconvolved drawdown pressure response and its Bourdet derivative on a diagnostic plot.

4.4 Production Surveillance Using Deep Learning Systematic Workflow

Data for this field case is from well 50H (in an unconventional reservoir) with both pressure and rate histories. To test the systematic workflow branching to the DeLSTM algorithm, we first hid 30 continuous hours of rate data and analyzed only rates displayed in solid blue lines in Fig. 90.

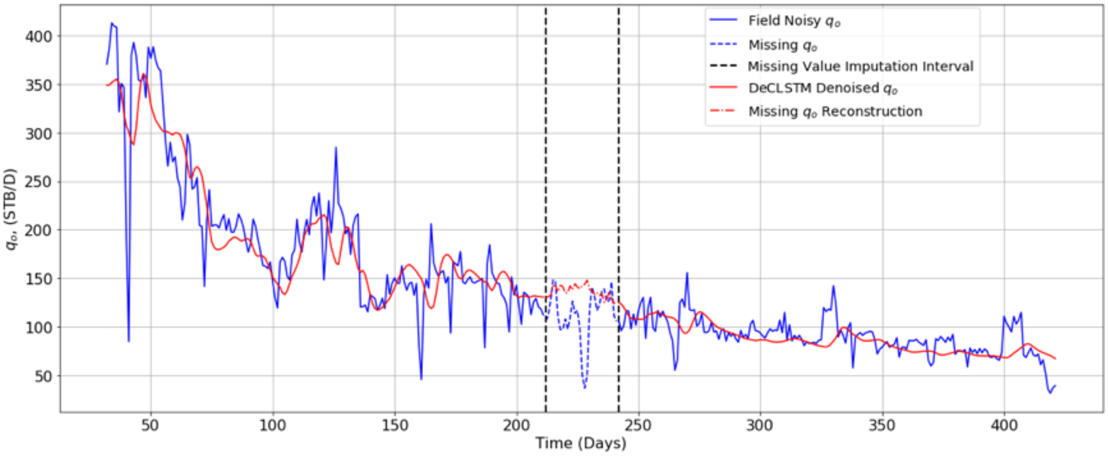


Fig. 90-DeLSTM denoising and missing-data imputation for well 50H.

Fig. 91 shows the results of DeLSTM denoising and missing interval imputation, and we observe a good alignment in overall tendencies. However, the model overestimates the imputated rates in the missing interval. Recall that DeLSTM is trained and utilized for prediction based only on the time dependence along the available sequence without introducing any physics. This could be the main reason that it limits the application of DeLSTM to a certain small period of missing-data imputation. Further investigation of this model limitation would be rewarding.

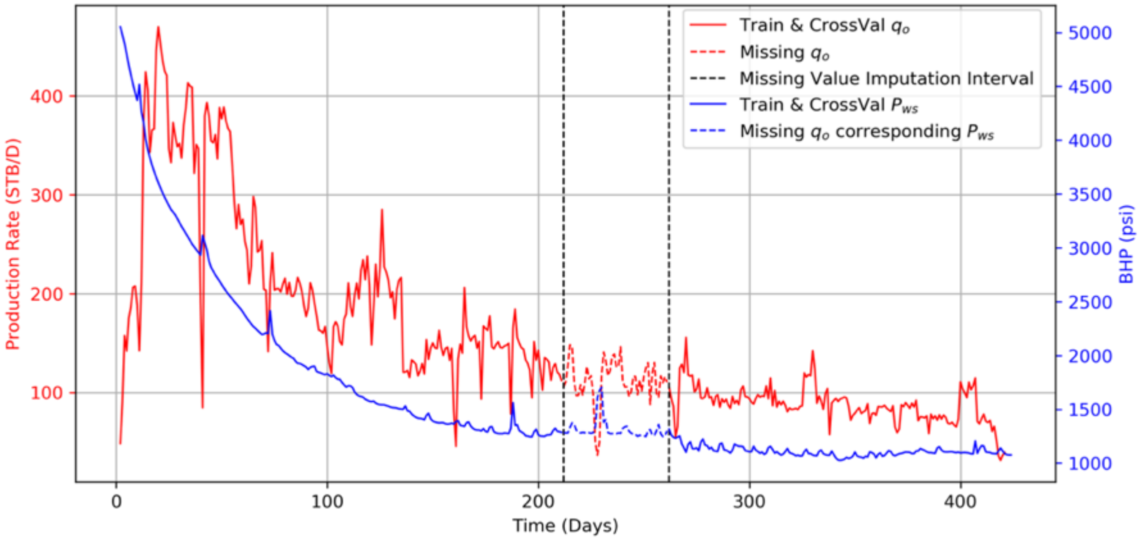


Fig. 91-Bottom hole pressure and flow rate histories with approximately 50 days of missing data for well 50H.

Once the pressure data is recorded along with the rate history, the SG-CLSTM can be used for denoising and missing-data imputation during production surveillance. Fig. 92 shows the rate changes along with the pressure changes, and, obviously, a level of noise exists in both data sets. Following the workflow in Fig. 19, the SG-denoised signals are illustrated in Fig. 92 and the reconstructed signals and missing interval imputation is compared with the SG denoised q_o shown in Fig. 93. By incorporating the pressure data, the reconstructed response for the entire surveyed period is improved. We attribute this result to the introduction of physics-based training features by honoring the transient flow mechanism. At this point, we further demonstrated the versatility of the systematic workflow and accuracy of missing-data regeneration using SG-CLSTM by surveying both pressure and rate data.

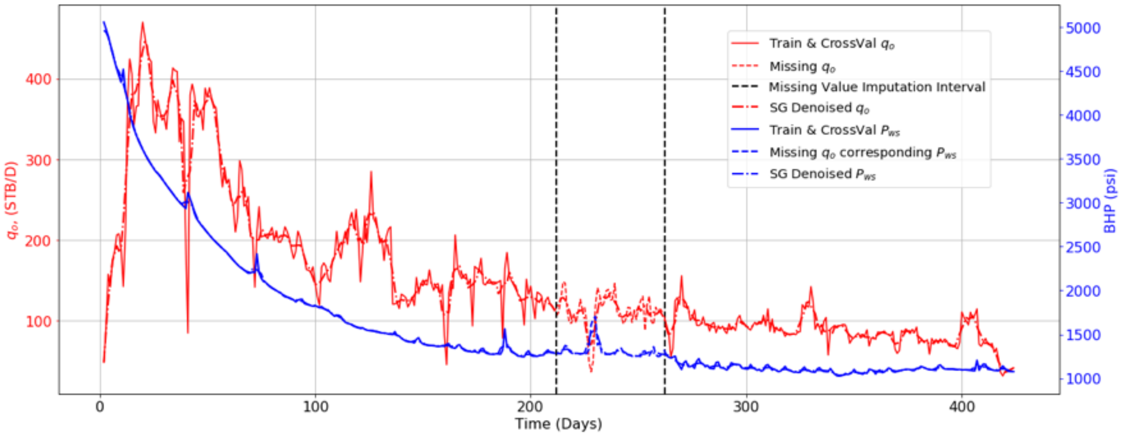


Fig. 92-SG smoothed BHP and production rate signals compared to original signals from well 50H.

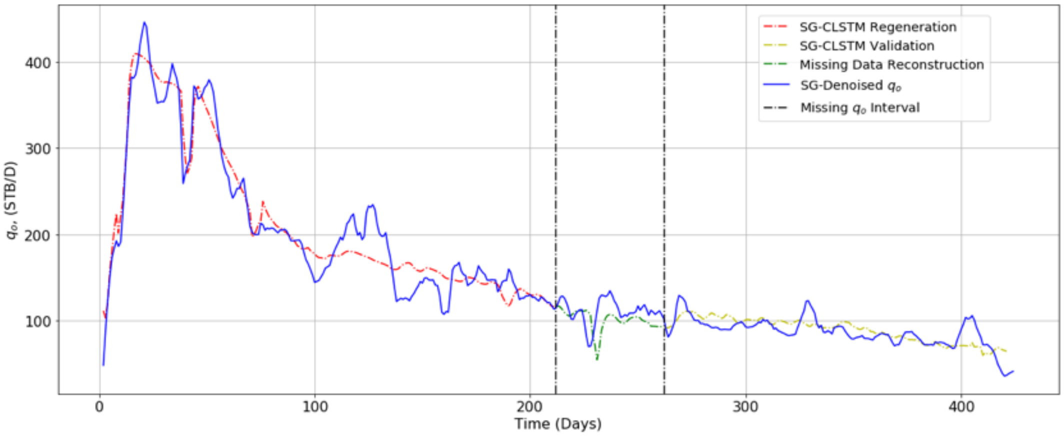


Fig. 93-Comparison of reconstructed and SG-denoised production rates for well 50H.

CHAPTER V

CONCLUSIONS

5.1 LEESN and BESN in Rate Transient Analysis

The objective of this algorithm presented in section 2.1 is to provide alternative approaches for rate/pressure deconvolution and to facilitate production forecasting. The physics-based training features and training algorithms play the key roles in successful application of deconvolution in production data analysis and production forecasting by honoring superposition transient flow physics. In the synthetic cases in section 3.1.6, sensitivity analysis based on the noise scale and training scale performed on LIESN and LEESN demonstrates LEESN's capability of handling a moderate noise level. An excessive level of noise might enlarge the deviation from the true flow regimes. Overall, we have provided a reliable diagnostic tool to use before forecasting production. Based on the accuracy provided by capturing correct physics in the proposed methodologies, we enhance the precision of production forecasting.

Part of Chapter V is reprinted with permission of the Society of Petroleum Engineers from Pan, Y., Zhou, P., Deng, L., Lee, J. 2019. Production Analysis and Forecasting for Unconventional Reservoirs Using Laplacian Echo-State Networks. Proceedings of SPE Western Regional Meeting, San Jose, California, 23-26 April. SPE-195243-MS. DOI:10.2118/195243-MS, and from Pan, Y., Deng, L., Lee, J. 2019. Data-Driven Deconvolution Using Echo-State Networks Enhances Production Data Analysis in Unconventional Reservoirs. Proceedings of SPE Eastern Regional Meeting, Charleston, West Virginia, 15-17 October. SPE-196598-MS. DOI: 10.2118/196598-MS, with permission conveyed through Copyright Clearance Center, Inc., and with permission of URTeC from Pan, Y., Bi, R., Zhou, P., Deng L., Lee, J. 2019. An Effective Physics-Based Deep Learning Model for Enhancing Production Surveillance and Analysis in Unconventional Reservoirs. Proceedings of Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July. URTEC-2019-145. DOI:10.15530/urtec-2019-145.Copyright 2019 URTeC.

Extra functionality for the probability of error interval can be provided by substituting the LEESN with BESN. This extra feature could critically increase confidence when forecasting future production. All in all, we demonstrated the capability of the proposed methodologies using a synthetic case and a field example. The workflow of LEESN regulates the consistency of how engineers implement the proposed algorithm from which the proposed approaches in production forecasting enhance the accuracy in reserves bookings. Moreover, as these algorithms are still a work in progress, further development and investigation are necessary to fully understand the limitations and capabilities.

5.2 LEESN in Pressure Transient Analysis

In section 4.2, we extended the Laplacian Eigenmaps coupled Echo State Network (LEESN) from deconvolution in rate transient analysis to pressure transient analysis and provide a more reliable deconvolution approach to facilitate the interpretation of pressure-rate data using the traditional diagnostic plot for unconventional reservoirs. The physics-based training features and training algorithms were reaffirmed regarding the critical aspect for successful application in deconvolution and production analysis. LEESN can cope with moderate noise and highly variable pressure-rate-time information accurately and rapidly because of its locality-preserving attribute. Using diagnostic tools including diagnostic plots or specialized plots based on the deconvolved response, we can easily and precisely recognize the flow patterns in addition to determining the corresponding reservoir properties. Finally, following the proposed workflow, we demonstrated the versatility and applicability of LEESN using synthetic and field cases.

5.3 Production Surveillance Using Deep Learning Approach

In section 3.3, we proposed a comprehensive and systematic workflow for production surveillance and to facilitate the production analysis and forecasting. With only the production rate being surveyed, DeLSTM can smooth the signals and impute missing data for a relatively short period. By surveying additional pressure histories along with rate data, the physics-based training features play key roles in successfully reconstructing the consecutive missing data and further assist deconvolution and forecasting. Consequently, we provided a reliable systematic workflow to apply prior to any end-user's analysis and demonstrated this with scenarios I and II cases and a field example. Automation can be achieved based on the accuracy of production profile completion, thereby enhancing the efficiency of data integration in production surveillance.

NOMENCLATURE

Roman symbols

A_{SRV}	The stimulated reservoir area, ft^2
a_t	Block activation input
$\overline{a}_t$	Block input
B	Formation volume factor, RB/STB
b	The vector form of $[b_a, b_i, b_f, b_o]$
b_a, b_i, b_f, b_o	Bias weights
c_t	Internal state
$\mathrm{c_t}$	Total compressibility psi-1
$\mathbf{D}$	Diagonal weight matrix where it can be determined using $D_{ii} = \sum_j W_{ji}$, the column sums of W, the weight matrix of the edge of adjacent matrix
d_{st}	Euclidean distance
E	Error Function
$\mathbf{F}$	The vector form of eigenvector
$\mathbf{f}, \mathbf{f_i}, \mathbf{f_j}$	The vector or matrix form of eigenvector
f_t	Forget gate

$\overline{f}_t$	Forget gate before activation function
G_t	Vector form of $[a_t, i_t, f_t, o_t]$
$h(x)$	tanh function
h	Reservoir height, ft
$\mathbf{I}$	Identity matrix or vector
i_t	Input gate
$\bar{i}_t$	Input gate before activation function
k	Reservoir matrix permeability, MD
L	Laplacian matrix, $L = \boldsymbol{D} - \boldsymbol{W}$
$\mathrm{L_{wb}}$	Length of horizontal wellbore, ft
M	The number of inputs
m	The slope of linear flow regime on specialist plot
N	The number of LSTM units
$\mathrm{N, N_y, N_t, N_r}$	Total number of observations, total number of output samples, total number of time series data, total number of reservoir nodes
n_f	The number of fractures
o_t	Output gate

$\overline{o}_t$	Output gate before activation function
p_i	Reservoir initial pressure, psi
p_{wf}, p_{ws}	Well bottomhole flowing pressure, Well bottom hole sand face pressure, psi
q, q_i	The surface flow rate, STB/D
s_w	Water saturation, %
t	Time, hrs or days
U	The vector form of $[U_a, U_i, U_f, U_o]$
$\mathrm{U_a}, \mathrm{U_i}, \mathrm{U_f}, \mathrm{U_o}$	Recurrent weights
$\mathbf{u}, \mathbf{u_i}, \mathbf{u}(t)$	The vector or matrix form of training inputs
$\mathbf{u}^\star$	The matrix form of prediction inputs under LEESN generative mode
V_{SRV}	The stimulated reservoir volume, ft^3
W	The vector form of $[W_a, W_i, W_f, W_o]$
W_a, W_i, W_f, W_o	Input weights
W_{ij}	Heat kernel weights for the adjacency matrix
$\mathbf{W_u}$	The matrix form of internal weights in reservoir temporal space
$\mathbf{W_x}, \mathbf{W_{\tilde{x}}}$	The vector form of input weights connecting input layer to reservoir temporal space

$\mathbf{W}_{\mathbf{y}}$	The vector form of output weights connecting reservoir temporal space to output layer
$\dot{\mathbf{W}}_{\mathbf{y}}$	Corresponding lower dimension output weights
$\mathrm{w}_{\mathrm{y}_{\mathrm{i}}}$	Output weights connecting reservoir temporal space to output layer
$\mathbf{X}_{\mathbf{i}}, \mathbf{X}_{\mathbf{j}}$	The matrix form of reservoir intermediate states
$\dot{\mathbf{X}}$	The lower dimension representation of reservoir states
$\widetilde{\mathbf{X}}$	The original representation of reservoir states after activation
$\widetilde{\mathbf{X}}^{\star}$	Reservoir original states after activation under model generative mode using new input data $\boldsymbol{u}^{\star}$
$\dot{\mathbf{X}}^{\star}$	Reservoir lower dimensional representatitve of original states corresponding to $\widetilde{\boldsymbol{X}}^{\star}$
$\mathbf{x}, \mathbf{x}_{\mathbf{i}}, \mathbf{x}(t)$	The vector or matrix form of reservoir intermediate states
$\tilde{\mathbf{x}}(t)$	The matrix form of reservoir intermediate states before leaky integrator unit activated
x_f	Fracture half-length, ft
x_t	Input vector at time t
$\mathbf{Y}$	The vector or matrix form of true target feature
$\widehat{\mathbf{Y}}$	The LEESN regeneration output
$\widehat{\mathbf{Y}}^{\star}$	Predictions using new input data $\boldsymbol{u}^{\star}$ under generative mode of LEESN
$\mathbf{y}, \mathbf{y}_{\mathbf{i}}, \mathbf{y}(t)$	The vector or matrix form of target outputs

$\hat{\mathbf{y}}, \hat{\mathbf{y}}_{\mathbf{i}}, \hat{\mathbf{y}}(t)$	The vector or matrix form of model prediction
y_t	Target output vector at time t
z^i	Sorted distance d_{st}

Greek letters

α	Leaking rate, the speed of the reservoir update dynamics discretized in time
α_i	The slope representing the regressed systems
β	The intercept representing the regressed systems
$\boldsymbol{\alpha}_{\boldsymbol{BR}}$	The hyper parameter governs the prior distribution over the output weights
β_r	Regularization coefficient of ridge regression
$\boldsymbol{\beta}_{\boldsymbol{BR}}$	Zero-mean Gaussian distribution
$\boldsymbol{\gamma}$	Intermediate parameter to update $\boldsymbol{\alpha}_{\boldsymbol{BR}}, \boldsymbol{\beta}_{\boldsymbol{BR}}$
$\boldsymbol{\epsilon}$	Residual between the prediction and target outputs
$\boldsymbol{\theta}^{\mathbf{T}}$	The vector of regression coefficients
λ, λ_i	Eigenvalues
λ_{BR}	Eigenvalues of the Hessian matrix of the error function in Bayesian ridge regression
μ	Fluid viscosity, cp

σ	Predefined heat kernel width
$\sigma(x)$	Sigmoid activation function
τ	Specified recursion depth, the recursion for practical reasons is needed to be cutoff at a specific point
φ	The activation function, tanh is selected in this paper
ϕ	Porosity, %
ϕ_{SRV}	Stimulated reservoir volume porosity
$\mathcal{H}$	The number of lower dimensional space representing the original reservoir space
κ_t	Kernel function can be expressed as $<,>$
κ	Intermediate parameter to update $\boldsymbol{\alpha}_{\boldsymbol{BR}}, \boldsymbol{\beta}_{\boldsymbol{BR}}$
$\boldsymbol{\Sigma}$	Intermediate parameter to update $\boldsymbol{\alpha}_{\boldsymbol{BR}}, \boldsymbol{\beta}_{\boldsymbol{BR}}$

REFERENCES

Ansley, C. F., and Kohn, R. (1984). On the estimation of ARIMA models with missing values. *Time series analysis of irregularly observed data* (pp. 9-37). Springer, New York, NY.

Antonelo, E., Camponogara, E., and Foss, B. 2017. Echo State Networks for Data-Driven Downhole Pressure Estimation in Gas-Lift Oil Wells. *Neural Networks*, 106-117. DOI:10.1016/j.neunet.2016.09.009.

Arps, J. J. 1945, Analysis of Decline Curves, *Transactions of the AIME*, 160, 228. SPE-945228-G, DOI:10.2118/945228-G.

Belkin, M. and Niyogi, P. 2003. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *Neural Computation*, 15, 1373-1396. DOI: 10.1162/089976603321780317.

Cao, W., Wang, D., Li, J., Zhou, H., Li, L., Li, Y. 2018. BRITS: Bidirectional Recurrent Imputation for Time Series. *Advances in Neural Information Processing Systems* (pp. 6775-6785).

Camilleri, L. 2010. Real-time flow rate obtained from ESP data, Hart Energy Publishing, Texas, USA.

Che, Z., Purushotham, S., Cho, K., Sontag, D., and Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific reports*, 8(1), 6085. DOI:10.1038/s41598-018-24271-9.

Dambre, J., Verstraeten, D., Schrauwen, B., and Massar, S. 2012. Information processing capacity of dynamical systems. *Scientific Reports*, 2, 512. DOI:10.1038/srep00514.

DeGolyer and MacNaughton, 2019. Well Performance in Unconventional Resources, Texas A&M Short Course, 02 March, 2019, College Station, Texas, USA.

Deng, L., Davani, E., Darabi, H., et al. 2018. Rapid and Comprehensive Artificial Lift Systems Performance Analysis Through Data Analytics, Diagnostics and Solution Evaluation. Proceedings of SPE Middle East Artificial Lift Conference and Exhibition, Manama, Bahrain, 28-29 November. SPE-192460-MS. Society of Petroleum Engineers. DOI: 10.2118/192460-MS.

Gers, F. A., Schmidhuber, J., and Cummins, F. (1999). Learning to forget: Continual prediction with LSTM. *in Proc. 9th int. Conf. Artif. Neural Network. (ICANN)*, (2), 850-855. DOI:10.1049/cp:19991218.

Gonzalez, S., Al Muhanna, T., Abdalla, W., Pandey, D., Al Naqi, A., Al Mezal, L., … Hamlaoui, A. (2018). Production Data Management Collaboration Effort in an Integrated Journey for More than 1,000 Wells in the Northern Kuwait Heavy Oil Fields. Proceedings of SPE International Heavy Oil Conference and Exhibition. SPE-193762-MS. DOI:10.2118/193762-MS.

Graves, A., and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5-6), 602-610. DOI:10.1016/j.neunet.2005.06.042.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10), 2222-2232. DOI:10.1109/TNNLS.2016.2582924.

Han, M. and Xu, M. 2018. Laplacian Echo State Network for Multivariate Time Series Prediction. *IEEE Transactions on Neural Networks and Learning Systems*, 29, 238-244. DOI:10.1109/TNNLS.2016.2574963.

Hermans, M. and Schrauwen, B. 2012. Recurrent Kernel Machines: Computing with Infinite Echo State Networks. *Neural Computation*, 24, 104-33. DOI:10.1162/NECO_a_00200.

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780. DOI:10.1162/neco.1997.9.8.1735.

Infosys, 2019, *Infosys production surveillance solution footprints*: http//:www.infosys.com/industries/oil-and-gas/industry-offerings/Pages/production-surveillance.aspx.

Jaeger, H. and Haas, H. 2004: Harnessing nonlinearity: Prediction chaotic systems and saving energy in wireless communication. *Science*, 304, 78-80. DOI:10.1126/science.1091277.

Jaeger, H., Lukoševičius, M., Popovici, D., and Siewert, U. 2007. Optimization and applications of echo state networks with leaky- integrator neurons. *Neural Networks*, 20, 335-252. DOI:10.1016/j.neunet.2007.04.016.

Kabir, C. S., and Lake, L. W. 2011. An Analytical Approach to Estimating EUR in Unconventional Reservoirs. Proceeding of SPE North American Unconventional Gas Conference and Exhibition, 14-16 June, The Woodlands, Texas, USA. SPE-144311-MS, DOI: 10.2118/144311-MS.

Kuchuk, F., Hollaender, F., Gok, I., and Onur, M. 2005. Decline Curves from Deconvolution of Pressure and Flow-Rate Measurements for Production Optimization and Prediction. Proceedings of SPE Annual Technical Conference and Exhibition, 9-12 October, Dallas, Texas, USA. SPE-96002-MS, DOI:10.2118/96002-MS.

Kuchuk, F. J., and Biryukov, D. 2012. Transient pressure test interpretation from continuously and discretely fractured reservoirs. Proceedings of SPE Annual Technical Conference and Exhibition, 8-10 October, San Antonio, Texas, USA. SPE-158096-MS, DOI:10.2118/158096-MS.

Kuchuk, F., Morton, K., and Biryukov, D. 2016. Rate-transient analysis for multistage fractured horizontal wells in conventional and un-conventional homogeneous and naturally fractured reservoirs. Proceedings of SPE Annual Technical Conference and Exhibition, 26-28 September, Dubai, UAE. SPE-181488-MS, DOI:10.2118/181488-MS.

Lee, J., Rollins, J. and Spivey, J. 2003. SPE Textbook Series, Volume 9 : Pressure Transient Testing. ISBN 978-1-55563-099-7.

Levitan, M. 2003. Practical Application of Pressure-Rate Deconvolution to Analysis of Real Well Tests. Proceedings of SPE Annual Technical Conference and Exhibition, 5-8 October, Denver, Colorado, USA. SPE-84290-MS, DOI:10.2118/84290-MS.

Li, D., Han, M., and Wang, J. 2012. Chaotic time series prediction based on a novel robust echo state network. *IEEE Transactions on Neural Networks and Learning Systems*, 23(5), 787-799. DOI: 10.1109/TNNLS.2012.2188414.

Li, G., Li, B. J., Yu, X. G., Cheng, and C. T. 2015. Echo state network with bayesian regularization for forecasting short-term power production of small hydropower plants. *Energies*, 8(10), 12228-12241. DOI: 10.3390/en81012228.

Liu, G. and Ehlig-Economides, C. 2018. Practical Considerations for Diagnostic Fracture Injection Test (DFIT) Analysis. *Journal of Petroleum Science and Engineering* **171**: 1133-1140. DOI: 10.1016/j.petrol.2018.08.035.

Liu, Y., and Horne, R. 2013. Interpreting Pressure and Flow-Rate Data From Permanent Downhole Gauges by Use of Data-Mining Approaches. *SPE Journal*, 18, 69-82. SPE-147298-PA. DOI:10.2118/147298-PA.

Lukoševičius, M. 2012. A Practical Guide to Applying Echo State Networks. Berlin, Germany, Springer.

Maass, W., Natschläger, T. and Markram, H. 2002. Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations. *Neural Computation*. 14, 2531-2560. DOI:10.1162/089976602760407955.

Mittal, R. (2018). Technology Focus: Production Monitoring/Surveillance (March 2018). *Journal of Petroleum Technology*, 70(03), 66-66. SPE-0318-0066-JPT. DOI:10.2118/0318-0066-JPT.

Olah, C., 2015, *Understanding LSTM Networks:* https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

Onur, M., and Kuchuk, F. J. 2012. A new deconvolution technique based on pressure-derivative data for pressure-transient-test interpretation. *SPE Journal*, 17, 307-320. SPE-134315-PA. DOI:10.2118/134315-PA.

Pathak, J., Hunt, B., Girvan, M., Lu, Z. and Ott, E. 2018. Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach. *Physical Review Letters*. 120, 024102. DOI:10.1103/PhysRevLett.120.024102.

Reynolds, A. C., Oliver, D. S., Li, G., Dong, Y., Zhao, Y., Eydinov， D., Gao, G., Chen, S., and Han, M., 2005. Data integration for the generation of high resolution reservoir models. Technical reports, Project No. DE-FC26-04NT15517, US DOE/NETL, Washington, DC(2005).

Savitzky, A., and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8), 1627-1639. DOI:10.1021/ac60214a047.

Scardapane, S. and Wang, D. 2017. Randomness in neural networks: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 7, DOI:10.1002/widm.1200.

Song, B. and Ehlig-Economides C A. 2011. Rate-normalized pressure analysis for determination of shale gas well performance. Proceedings of North American Unconventional Gas Conference and Exhibition, 14-16 June, The Woodlands, Texas, USA. SPE-144031-MS, DOI:10.2118/144031-MS.

Sun, Q. and Ertekin, T. 2015. The Development of Artificial-neural-network-based Universal Proxies to Study Steam Assisted Gravity Drainage(SAGD) and Cyclic

Staem Stimulation (CSS) Processes. SPE Western Regional Meeting, 27-30 April, Garden Grove, California, USA. SPE-174074-MS, DOI:10.2118/174074-MS.

Sun, Q. and Ertekin, T. 2017. Structuring an artificial intelligence based decision making tool for cyclic steam stimulation processes. *Journal of Petroleum Science and Engineering*, 154, 564-575. DOI:10.1016/j.petrol.2016.10.042.

Tian, C. and Horne, R. 2015. Applying Machine Learning Techniques to Interpret Flow Rate, Pressure and Temperature Data From Permanent Downhole Gauges. Proceedings of SPE Western Regional Meeting, 27-30 April, Garden Grove, California, USA. SPE-174034-MS, DOI:10.2118/174034-MS.

Tian, C. and Horne, R. 2019. Applying Machine Learning Techniques to Interpret Flow Rate, Pressure and Temperature Data From Permanent Downhole Gauges. *SPE Reservoir Evaluation & Engineering*, 386-401. SPE-174034-PA, DOI:10.2118/174034-PA.

Udegbe, E., Morgan, E., Srinivasan, S. (2019). Big-Data Analytics for Production-Data Classification Using Feature Detection: Application to Restimulation-Candidate Selection. *SPE Reservoir Evaluation & Engineering*, SPE-187328-PA. DOI:10.2118/187328-PA.

Valko, P. and Lee, W. 2010. A Better Way To Forecast Production From Unconventional Gas Wells. Proceedings of SPE Annual Technical Conference and Exhibition, 19-22 September, Florence, Italy. SPE-134231-MS, DOI:10.2118/134231-MS.

von Schroeter, T., Hollaender, F., and Gringarten, A. 2001. Deconvolution of Well Test Data as A Nonlinear Total Least Squares Problem. Proceedings of SPE Annual Technical Conference and Exhibition. 30 September – 3 October, New Orleans, Louisiana, USA. SPE-71574-MS, DOI:10.2118/71574-MS.

von Schroeter, T., Hollaender, F. and Gringarten, A. 2004. Deconvolution of Well-Test Data as a Nonlinear Total Least-Squares Problem. *SPE Journal*, 9,375-390. SPE-77688-PA, DOI:10.2118/77688-PA.

Wang, W., Su, Y., Sheng, G., Cossio, M., and Shang, Y. (2015). A mathematical model considering complex fractures and fractal flow for pressure transient analysis of fractured horizontal wells in unconventional reservoirs. *Journal of Natural Gas Science and Engineering*, 23, 139-147. DOI: 10.1016/j.jngse.2014.12.011.

Wu, J., Huang, D. Y., Xie, L., and Li, H. (2017, August). Denoising Recurrent Neural Network for Deep Bidirectional LSTM Based Voice Conversion. *INTERSPEECH* (pp. 3379-3383). DOI:10.21437/Interspeech.2017-694.

Yang, Q., Jin, B., Banerjee, D., and Nasrabadi, H. (2019). Direct visualization and molecular simulation of dewpoint pressure of a confined fluid in sub-10 nm slit pores. *Fuel*, 235, 1216-1223. DOI:10.1016/j.fuel.2018.08.050.

Zhang, D., Yuntian, C., Jin, M. (2018). Synthetic well logs generation via Recurrent Neural Networks. *PETROLEUM EXPLORATION AND DEVELOPMENT-ELSEVIER-*, 45(4), 629-639.

APPENDIX A

TRANSIENT FLOW EQUATION FOR GAS

The equations in Table 3 and Table 7 can be applied only to slightly compressible fluids, but when pseudo pressure is substituted for pressure and the pseudo time is substituted for time, both tables can be applied to gases.

The pseudo pressure is defined as:

$$P_p = 2\int_{p_0}^{p} \frac{p}{\mu z}\, dp \qquad (67)$$

where P_p is pseudo pressure in $\mathrm{psi}^2/\mathrm{cp}$, and the integral group $\frac{p}{\mu z}$ requires numerical integration.

Pseudo time is defined and expressed as:

$$t_p = \mu_i \mathrm{c_{ti}} \int_{0}^{t} \frac{dt}{\bar{\mu}\bar{\mathrm{c}}_{\mathrm{t}}} \qquad (68)$$

where μ_i is the initial gas viscosity in cp, $\mathrm{c_{ti}}$ is the initial total compressibility in psi^{-1}, $\bar{\mu}$ is the viscosity based on average reservoir pressure in cp, $\bar{\mathrm{c}}_{\mathrm{t}}$ is the total compressibility based on average reservoir pressure in psi^{-1}.