Physics-informed Machine Learning for Real-time Unconventional Reservoir Management

Maruti K. Mudunuru, Daniel O'Malley, Shriram Srinivasan, Jeffrey D. Hyman, Matthew R. Sweeney, Luke Frash, Bill Carey, Michael R. Gross, Nathan J. Welch, Satish Karra, Velimir V. Vesselinov, Qinjun Kang, Hongwu Xu, Rajesh J. Pawar, Tim Carr, Liwei Li, George D. Guthrie, Hari S. Viswanathan

Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM-87544.
Department of Geology & Geography, West Virginia University, Morgantown, WV 26506.
{maruti,omalled,shrirams,jhyman,msweeney2796,lfrash,bcarey,michael_gross,nwelch,satkarra,vvv,qkang,hxu,rajesh,geo,viswana}@lanl.gov, {tim.carr,liwei.li}@mail.wvu.edu

Abstract

We present a physics-informed machine learning (PIML) workflow for real-time unconventional reservoir management. Reduced-order physics and high-fidelity physics model simulations, lab-scale and sparse field-scale data, and machine learning (ML) models are developed and combined for real-time forecasting through this PIML workflow. These forecasts include total cumulative production (e.g., gas, water), production rate, stage-specific production, and spatial evolution of quantities of interest (e.g., residual gas, reservoir pressure, temperature, stress fields). The proposed PIML workflow consists of three key ingredients: (1) site behavior libraries based on fast and accurate physics, (2) ML-based inverse models to refine key site parameters, and (3) a fast forward model that combines physical models and ML to forecast production and reservoir conditions. First, synthetic production data from multi-fidelity physics models are integrated to develop the site behavior library. Second, ML-based inverse models are developed to infer site conditions and enable the forecasting of production behavior. Our preliminary results show that the ML-models developed based on PIML workflow have good quantitative predictions (>90% based on R^2-score). In terms of computational cost, the proposed ML-models are $\mathcal{O}(10^4)$ to $\mathcal{O}(10^7)$ times faster than running a high-fidelity physics model simulation for evaluating the quantities of interest (e.g., gas production). This low computational cost makes the proposed ML-models attractive for real-time history matching and forecasting at shale-gas sites (e.g., MSEEL – Marcellus Shale Energy and Environmental Laboratory) as they are significantly faster yet provide accurate predictions.

1. Introduction

Energy extraction from conventional resources involves producing crude oil, natural gas, and its condensates from rock formations that have high porosity and permeability (Bahadori, 2017). These rock formations are found below an impermeable rock. However, energy extraction from unconventional hydrocarbon resources (Ahmmed and Meehan, 2016) involves using advanced drilling and stimulation techniques (e.g., long horizontal laterals and multi-stage hydraulic fracturing) to extract crude oil and natural gas that are trapped in the pores of relatively impermeable sedimentary rocks (e.g., shale, tight sandstones).

Typically, unconventional reservoirs have porosity in the range of 0.04-0.08 and matrix permeability on the order of nanodarcies (10^{-16}-10^{-20} m^2) (Rezaee, 2015; Belyadi et al., 2019). Instead of the porous flow that dominates conventional reservoirs, fracture flow dominates unconventional reservoirs, with natural fractures dissecting the matrix and intersecting with the hydraulic fractures. As result, energy extraction is more difficult than conventional reservoirs. Model-based optimization of unconventional reservoirs is also challenging because due to the long horizontal laterals there is insufficient site data to inform high-fidelity physics models (Mohaghegn, 2017; Belyadi et al., 2019). Despite these challenges and due to the abundance of unconventional resources, with reserves projected to last for many decades, energy extraction from these resources have gained prominence in recent years (Briefing, 2013 and Weijermars, 2014). Current extraction efficiency from unconventional reservoir is very low (~5-10%) for tight oil and ~20% for shale gas (Sandrea, 2007; Muggeridge et al., 2014) compared to conventional reservoirs (~20-40%) (Zitha et al., 2008). This is because the impact of resource development processes (e.g., slow drawdown or fast drawdown) and underlying physics that determine the energy extraction from the impervious rocks are poorly understood (Rezaee, 2015; Belyadi et. Al., 2019). State-of-the-art workflows for unconventional reservoir management are data-driven, which perform poorly beyond their training regimes. Hence, innovative extraction strategies (e.g., pressure-drawdown management) coupled with advanced workflows (e.g., physics-informed machine learning) are needed to improve the hydrocarbon recovery efficiency (Seales et al., 2017; Lougheed et al., 2017; Mirani et al., 2018). *In this paper, we present a physics-informed machine learning (PIML) workflow (Fig.1) to address unconventional production for real-time reservoir management. One of the goals of this PIML workflow (Fig.2) is to develop fast and accurate ML-models grounded in physics for real-time history matching and production forecasting in a fracture shale gas reservoir.*

1.1 State-of-the-art Workflows and Key Gaps

Current workflows for unconventional reservoirs are predominantly based on production decline curve analysis and its extensions (Wu et al., 2013; Sun 2015), data-driven machine learning (ML) approaches (Holdaway 2014; Chaki 2015; Puyang et al., 2015; Carvajal et al. 2017; Mohaghegn, 2017), and/or extension of physics-based conventional reservoir workflows (Rezaee, 2015; Rajput and Thakur, 2016; Belyadi et al., 2019). Decline curve analysis provides empirical models to forecast production data based on the past production history. This type of approach lacks physics and in-depth knowledge of the site behavior is not included in the forecasting models. The data-driven ML approaches perform poorly when faced with uncertain, missing, and sparse data – common problems with existing datasets related to unconventional reservoirs. Moreover, the data-driven ML analyses perform poorly in making forecasts outside of their training regimes, and the exploration of novel production strategies fundamentally requires extrapolation (where ML struggles) as opposed to interpolation (where ML excels). The physics-based workflows adopted for modeling conventional reservoirs use extensively available site-characterization data (which is acquired over months or years) for history-matching. Even though large unconventional reservoir data (e.g, fiber-optics) are sampled at sparse locations, simply combining all the data together would not improve the accuracy of real-time forecasting. This is because reservoir conditions change considerably from one basin to another basin. Therefore, physical constraints need to be incorporated in workflows. Existing workflows employ high-fidelity physics models to perform simulations, which are expensive to run. For example, it takes several days to months to run reservoir-scale model simulations (Rezaee, 2015; Rajput and Thakur, 2016; Belyadi et al., 2019) with degrees-of-freedom in the $\mathcal{O}(10^8)$ on state-of-the-art HPC machines. As a result, these conventional reservoir workflows are not ideal for usage in comprehensive uncertainty quantification studies, which require 1000s of forward model runs. To overcome the key gaps associated with existing workflows, we propose a PIML workflow (Fig.1 and Fig.2) to accelerate the development of ML-models while constraining it with physics. The aim is (1) to develop a library from a combination of site observations and physics-based models that is representative of unconventional reservoir site behavior, and (2) to develop fast and accurate ML-models for real-time history matching to refine key site parameters and forecast production quantities of interest (QoIs) with uncertainty estimates. Key production QoIs include, total cumulative production of hydrocarbons, gas and water production rates, stage-specific production, and spatial evolution of residual gas, reservoir pressure, temperature, and stress fields based on a user-defined pressure-drawdown strategy.

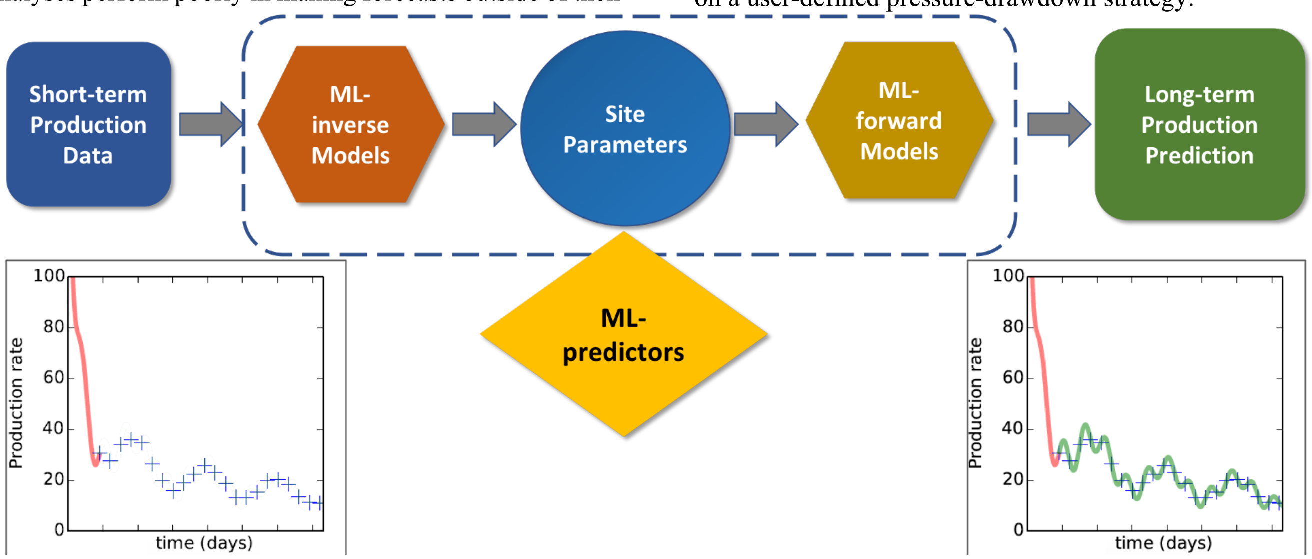

Figure-1: Physics-informed machine learning (PIML) workflow for reservoir management.

2. Physics-informed Machine Learning

2.1 Innovation and proposed approach

The proposed approach to develop the PIML workflow consists of three key steps: (Step-1) Development of site behavior libraries based on fast and accurate physics, (Step-2) Development of ML-based inverse models to infer key site parameters, and (Step-3) Development of fast forward models that combine physical models with ML to forecast production and reservoir conditions.

PIML Workflow Step-1: Development of a site behavior library involves generating synthetic data for a range of possible site characteristics. This includes a large number of runs from a fast physics-based reduced-order model and a smaller number of runs from a high-fidelity, full physics model. The fast physics models (e.g., Patzek models) allow us to quickly model and build a site library on the evolving reservoir data. These models allow us to efficiently explore the parameter space. Moreover, combining fast physics models with ML allows us to identify the important physical processes or dominant mechanisms that must be represented during full physics simulations. The dominant mechanisms at different stages of production include

- **Early stages of production (<1 year):** Primary fracture creation, geometry, and network connectivity, primary fracture behavior, propped fracture behavior, anisotropic permeability of fractures.
- **Middle stages of production (~1-5 years):** Secondary fractures and their permeabilities, shear fracture geometries, and geochemical impacts of hydraulic fluids and formation water.
- **Late stages of production (~5-10 years):** Matrix porosity and transport properties, water imbibition impacts, adsorbed gas properties, pore structure distribution.

These mechanisms are not accounted for in the fast physics models but are needed to describe the reservoir behavior at different stages of gas and water production. However, simulating these mechanisms using full physics models for entire parameter space is computationally intractable. *As a result, a fast physics model library along with ML are used to guide and improve the development of the full physics model library (which consist of complex 3D simulations of matrix-fracture interactions) for characterizing site behavior.*

PIML Workflow Step-2: The ML-inverse model is developed on this site behavior library using a transfer learning approach (Pan et al., 2009; Goodfellow et al., 2016; Yamada et al., 2018) where the numerous runs from the reduced order model are used to train an initial inverse model, then the smaller number of runs from the high-fidelity model are used to fine-tune the inverse model. This effectively represents a multi-fidelity approach to training the ML inverse model. This ML-inverse model provides capabilities for real-time history matching to update the key parameters (e.g., rock permeability, rock porosity, gas transport properties). Sensitivity analysis (e.g., Sobel indices, Random Forests) is performed to provide quantitative information on the key sensitive parameters (e.g, matrix permeability and porosity, fracture network parameters) that influence shale gas production rates.

PIML Workflow Step-3: The physics-based reduced order forward model uses these calibrated parameters for real-time forecasting. This model is trained on the site libraries along with evolving production data. Moreover, it allows for various operational decisions (e.g, slow drawdown vs. fast drawdown) to be evaluated relative to future outcomes. The ML-inverse and physics-based reduced order forward model can be combined to provide uncertainty estimates on the production quantities of interest (e.g., remaining hydrocarbon-in-place, spatial evolution of pressure and temperature, shale gas and water production as a function of time).

2.2 Key Challenges

The key challenges to accelerate the proposed PIML workflow include:

1Q. How do we enhance the information content and fill the gaps in the limited unconventional site-data for PIML analyses? Specifically, how to use the short-time gas production data (e.g., 30-120 days) to forecast the long-term performance (e.g., 1-5 years) of an unconventional shale gas well?

2Q. What is the minimum number of high-fidelity physics model simulations (e.g., PFLOTRAN, FEHM, dfnWorks) needed to develop a gas production library that is representative of shale gas sites behavior? How do we improve the full physics models to accurately represent the site behavior?

3Q. How can we use the information learned from reduced-order physics models (e.g., Patzek model, graph-based models) to inform high-fidelity physics model simulations?

4Q. What are the key sensitive parameters (e.g., matrix and fracture properties) in high-fidelity physics models that influence short-term and long-term gas production rates?

5Q. What are the key operational parameters (e.g., pressure drawdown strategies) that can be used to inform decisions in real-time, leading to optimized production?

2.3 Our Hypothesis

Our hypotheses/approaches to address the key challenges include:

1A. Augment limited unconventional reservoir data (e.g., short-term production data, well logs) with lab-scale experimental data, reduced-order physics simulation data, and high-fidelity physics simulation data. Short-term production data (e.g., 30-120 days) contains statistical information (e.g., matrix and fracture network properties) that can help us predict the long-term production data (e.g., 1-5 years).

2A. Reduced-order physics models provide information on the key sensitive parameters needed to inform high-fidelity physics model simulations. As we obtain more production data and/or site data, a ML-based active-feedback loop is used to improve full physics models. In this active feedback loop, ML-inverse model is used to update and constrain the key parameters based on newly available reservoir data. Based on these updated key parameters, site-behavior libraries are also updated to account for new production data.

3A. Transfer learning can be used to provide link and transfer information from reduced-order physics to high-fidelity physics.

4A. Sensitivity analysis (e.g., Sobel indices, Random Forests) can provide quantitative information on the key sensitive parameters (e.g, matrix permeability, matrix lithology, fracture length and orientation, stage spacing, hydrocarbon-in-place) that influence short-term and long-term gas production rates through feature importance.

5A. Maximizing early production may not maximize total recovery efficiency. Optimal pressure management strategies are needed to enhance recovery efficiency. One of our hypotheses is that slower drawdown rates can lead to improved recovery efficiency in long-term shale gas production (e.g., 5-10 years).

2.4 Details of the Proposed Approach

Fig.1 shows the overall PIML workflow for reservoir management. Short-time production data is fed to ML-inverse model to perform history-matching and infer key site parameters. These key parameters are then fed to ML-forward model to forecast long-term production QoIs. Fig.2 and Fig.3 provide more details on our PIML workflow. These figures show how the site behavior library is constructed and used to develop ML-models for history matching and real-time forecasting. Physically-realistic synthetic data is generated for a range of possible site characteristics using multi-fidelity physics models. This synthetic data provides insights on relevant features on poorly constrained site parameters and production scenarios that are not-yet observed. The site behavior library can be updated actively as the field-scale data becomes available overtime. This allows us to update or prune parameter combinations that are inconsistent with site characteristics. Note that any simulation platform (e.g., ECLIPSE, INTERSECT, tNavigator, CMG suite, Landmark Nexus, MRST, BOAST, OPM) (see ref. PetroMehras) can be used to generate synthetic data for the site behavior library.

The reduced-order physics models (Patzek et al., 2013) to generate synthetic data are given by

$$\frac{\partial \tilde{m}}{\partial \tilde{t}} - \frac{\partial}{\partial \tilde{x}}\left(\frac{\alpha(\tilde{m})}{\alpha_i}\frac{\partial \tilde{m}}{\partial \tilde{x}}\right) = 0 \tag{1}$$

where $\tilde{t}$ is the dimensionless time, $\tilde{x}$ is the dimensionless distance, and $\tilde{m}$ is the real gas pseudopressure. Eq.(1) corresponds to reduced-order physics of gas flow in fractured porous rock. These are given as follows

$$\tilde{t} = \frac{t}{\tau} \qquad \tau = \frac{d^2}{\alpha_i} \qquad \tilde{x} = \frac{x}{d} \qquad \alpha_i = \left.\frac{k}{\phi s_g \mu_g c_g}\right|_{Initial\ p,T}$$

$$\tilde{m} = \frac{1}{2}\left(\frac{[c_g p]\mu_g Z_g}{p^2}\right)_i m(x,t)$$

$$m(p,x,t) = 2\int_{p_{bhp}}^{p} \frac{p\,dp}{\mu_g Z_g(p,T)}$$

$$\alpha(m(p)) = \left.\frac{k(p)}{[\phi s_g + (1-\phi)\mathcal{K}_a\,]\mu_g c_g}\right|_{Evolving\ p,T} \tag{2}$$

where τ is the characteristic inference time, d is the half-distance between the hydrofractures, α_i is the initial hydraulic diffusivity, k is the fractured rock effective permeability dependent on reservoir pressure, ϕ is the fractured rock effective porosity, s_g is the fraction of pore space occupied by the gas, μ_g is the gas viscosity, c_g is the isothermal compressibility of gas, $\mathcal{K}_a$ is the differential equilibrium partitioning coefficient of gas, and Z_g is the compressibility factor of gas. These gas properties are dependent on evolving reservoir pressure and temperature. Eq. (1) is a nonlinear pressure diffusion equation, which is solved numerically. The cumulative production of gas mass is given as follows

$$\mathfrak{M}(\tilde{t}) = \mathcal{M}\int_0^{\tilde{t}} \left.\frac{\partial \tilde{m}}{\partial \tilde{x}}\right|_{\tilde{x}=0} dt' \tag{3}$$

where $\mathcal{M}$ is the initial hydrocarbon-in-place.

The expensive full physics models to simulate gas flow and transport in fractured porous media (Rezaee, 2015; Salama et al., 2017; Belyadi et al., 2019) are given by

$$\frac{\partial \phi s_g \rho_g}{\partial t} + \nabla\cdot\rho_g \boldsymbol{q} = 0 \tag{4}$$

$$\boldsymbol{q} = -\frac{k(p)}{\mu_g}\nabla(p - \rho_g g z) \tag{5}$$

$$\frac{\partial \phi c}{\partial t} + \nabla\cdot(c\boldsymbol{q} - \phi s \lambda D \nabla c) = 0 \tag{6}$$

where ρ_g is the density of the gas which is dependent on the reservoir pressure, $\boldsymbol{q}$ is the Darcy flux, c is the gas concentration, λ is the tortuosity, and D is the fracture rock effective diffusivity. Eq.(4) and Eq.(5) model the gas flow under varying reservoir and bottom hole pressures. The underlying assumptions include Darcy's flow and Fick's law. Adsorption and non-pore refinement effects on phase behavior are ignored. Eq.(6) models the gas transport from fractured stage to horizontal well based on the initial hydrocarbon-in-place. The amount of gas extracted from the pair of hydrofractures at a given section of the well is equal to $c\boldsymbol{q}$. Different types of equation of state (EOS) models are used to evaluate the gas density. These include ideal gas law, exponential pressure-dependent model, and Redlich-Kwong-Soave model. The corresponding EOS model expressions are given by

$$\rho_g = \frac{pM_g}{RT}$$

$$\rho_g = \rho_0 e^{\beta\times(p-p_0)}$$

$$\rho_g = \frac{pM_g}{Z_g(p,T)RT} \tag{7}$$

Note that it is possible to incorporate nanopore confinement effects (e.g,shifts in bubble or dew points) into EOS models to account for density changes. For example, see ref. Islam et al., 2015, Tan and Piri, 2015, and Liu and Zhang, 2019.

Through these multi-fidelity physics models, the reservoir physical behavior is captured accurately. Gas flow and transport mechanisms are accounted through conservation of mass and equation of state for real gases. State-of-the-art simulators (e.g., PFLOTRAN, dfnWorks) are used to develop high-fidelity simulation data. These simulators use finite volume methods and Newton-Raphson method to solve the discretized system of nonlinear equations given by Eq.(4)-(7). Moreover, these simulators account for accurate meshing of fractures, matrix, and upscaling of fracture network properties for reservoir-scale high-fidelity physics simulations.

Fig.3 shows the PIML workflow to create efficient inverse models to infer key site parameters from production information. First, we sample the relevant regions in parameter space. Two site behavior libraries are developed based on this sampling. One comes from running the reduced-order physics model with a large set of samples and the other comes from running the full physics model with a smaller set of samples. The first library (based on the reduced-order physics model) is used to train an initial ML-inverse model. The ML-inverse model is then fine-tuned with the library from the high-fidelity physics model using transfer learning, producing a final ML-inverse model. During this fine-tuning process the weights of the neural network are fine-tuned. The ML-inverse model takes past production as input and produces physical parameters as output. These physical parameters can then be fed into the reduced-order physics model. The loss function for this ML-inverse model is defined in terms of how well it works in combination with the reduced-order physics model at predicting future production. Note that the second library can also be augmented with field production data to improve the realism of the ML-inverse model. This fine-tuning represents a multi-fidelity approach to machine learning where a large dataset is generated with a reduced-order physics model and a smaller dataset is generated with a high-fidelity, expensive full physics model. This multi-fidelity approach allows us to perform real-time history matching on new production data to determine critical site parameters that can be used to accurately predict future production.

3. Results

Fig.4 shows a DFN model of a single stage based on field data from the Marcellus Shale Energy and Environment Laboratory (MSEEL) shale gas site. This model was built using dfnWorks, which is a computationally expensive full physics software suite used to generate high-resolution representations of DFNs (Hyman et al., 2015). This high-fidelity meshing of the fracture network is critical to accurately capture the physical processes in a fractured shale gas reservoir. To capture the matrix effects, we generate an octree-refined continuum mesh (grey color in Fig.4) based on the DFN. The original DFN model in Fig.4 consists of three hydraulic fractures and a swarm of natural fractures that are connected to hydraulic fractures. While generating the continuum mesh, the DFN model is simultaneously upscaled to account for matrix-fracture interactions, which results in accurate permeability and porosity values that are needed to simulate gas flow and transport in fractured shale. The final mesh contains approximately 500,000 mesh cells.

Fig.5 shows the flow and transport simulation using PFLOTRAN with a Barton-Bandis stress relationship (Barton and Bandis, 1990). This figure shows the drainage over a period of 10 years. For gas flow simulations (left), the well is maintained at 12MPa and the reservoir initial pressure is at 21MPa. For gas transport simulations (right), the initial hydrocarbon-in-place is assumed to spread in the entire fracture stage and the figure shows the transport of hydrocarbon to the well over time at two different vertical heights. The main inference from these simulations is that the characteristic behavior of drainage is tied to hydraulic fractures. Our future work involves accounting for heterogeneity of hydrocarbon distribution in the matrix for production forecasts.

Fig.6 shows encouraging results of long-term production forecasts using ML-forward models. The red color represents the short-term production, which is used along with the site behavior library built on reduced-order physics models to infer key site parameters (e.g., initial hydrocarbon-in-place, hydraulic diffusivity). History-matching is performed on the short-term production data. That is, the ML-inverse model is trained on 90 days of production data (red color). Then, the resulting ML-model with site behavior library based on reduced-order physics is used to predict future production (green color). The markers represent the long-term production data and the solid green color line represents the ML-model forecasts. Quantitatively, the prediction accuracy based on R^2-score is > 90%. From this figure, it is clear that ML-model predictions, which are combined with physics are able to accurately represent the long-term production data.

3.1 Discussion

We note that if another site/formation (e.g., Woodford, Haynesville, Fayetteville, Barnett, Utica, EagleFord) has a similar parameter range, we can use a set of ML techniques derived from transfer learning to model this site/formation. Transfer learning helps us to transfer knowledge gained from one site (e.g., Marcellus) to another site (e.g., Woodford, Barnett). But the developed ML-models for one site (e.g., MSEEL) need fine-tuning (or minimal retraining the neural networks) to transfer knowledge across shale sites/formations. This is not burdensome when compared to developing a new site behavior library and ML-model for a different site altogether. The transfer learning approach is attractive for tasks where reusability of ML-models for similar types is of great importance.

If the production curve contains high frequency content or oscillations, which might need addressing, there are various ways to incorporate this in the ML analyses and physics-based models. For example, from the production curve and bottom hole pressure, we can extract dominant frequencies or a range of frequencies we are interested in modeling through Fast Fourier Transformation (FFT). This FFT transformation of bottom hole pressure and production curve vs. time provides in-depth information on quantitative aspects of high-frequency oscillations to be incorporated in physics models (e.g., $p_{bhp} = a_1 \sin \omega_1 t + a_2 \sin \omega_2 t + \dots + a_{n-1} \sin \omega_{n-1} t + a_n \sin \omega_n t$) when developing the site behavior library and pressure management strategies (e.g., drawdown frequencies).

4. Conclusions

In this paper, we have presented a PIML workflow for real-time history matching and forecasting of gas production QoIs. The workflow coupled the strengths of machine learning with the predictability of physics-based models for real-time history-matching and forecasting. The PIML workflow used short-term production data and site behavior libraries to perform real-time history matching. Site behavior libraries are developed based on many runs of a reduced-order physics models and a smaller number of runs of expensive full physics models. The initial ML-inverse model trained on the reduced physics site library and short-term production data provided us key site parameters, which are hydraulic diffusivity and initial hydrocarbon-in-place. This initial ML-inverse model is then fine-tuned with the library from the expensive full physics models using transfer learning. The expensive full physics model simulations were developed using dfnWorks and PFLOTRAN simulators. These high-fidelity simulations account for matrix-fracture interaction, which is needed to accurately simulate gas flow and transport in fractured shale reservoirs. Moreover, from these simulations we inferred that the characteristic behavior of drainage is tied to hydraulic fractures. This high-fidelity simulation library was used to fine-tune the ML-inverse model. Our ongoing work seeks to advance and test the PIML workflow, site behavior libraries, and ML-models for pressure management (e.g., slow drawdown vs. fast drawdown) to optimize recovery at MSEEL. Our preliminary results shown in this paper are encouraging for use of site behavior libraries with ML-inverse model to address this problem.

Acknowledgements

This work was funded by the U.S. Department of Energy (DOE) – Office of Fossil Energy (FE) through its Oil and Natural Gas Research Storage Program as implemented by the National Energy Technology Laboratory. We would specifically like to thank Jared Ciferno for his support and guidance. LANL is operated by Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001). The opinions expressed in this paper are those of the authors and do not necessarily reflect that of the sponsors.

References

Ahmmed, U. and Meehan, D. N., (2016), *Unconventional Oil and Gas Resources: Exploitation and Development*, CRC Press, Boca Raton, FL, USA.

Bahadori, A. (2017), *Fluid Phase Behavior for Conventional and Unconventional Oil and Gas Reservoirs*, Elsevier, Oxford, UK.

Barton, N., and Bandis, S. (1990), Review of predictive capabilities of JRC-JCS model in engineering practice. *In Rock Joints, Proc. Int. Symp on Rock Joints*, Loen, Norway (eds N. Barton and O. Stephenson) (pp. 603-610).

Belyadi, H., Fathi, E., and Belyadi, F., (2019), *Hydraulic Fracturing in Unconventional Reservoirs: Theories, Operations, and Economic Analysis*, Gulf Professional Publishing, Elsevier, Cambridge, Massachusetts, USA.

Briefing, U. S. S. (2013), International energy outlook 2013, US Energy Information Administration.

Carvajal, G., Maucec, M., and Cullick, S. (2017), Intelligent Digital Oil and Gas Fields: Concepts, Collaboration, and Right-time Decisions. Gulf Professional Publishing, Elsevier, Cambridge, Massachusetts, USA.

Chaki, S., (2015), Reservoir Characterization: A Machine Learning Approach. *arXiv preprint arXiv:1506.05070*.

Goodfellow, I., Bengio, Y., and Courville, A. (2016), *Deep learning*. MIT press.

Holdaway, K. R. (2014), *Harness Oil and Gas Big Data with Analytics: Optimize Exploration and Production with Data-driven Models*. John Wiley & Sons, New Jeresey, USA.

Hyman, J. D., Karra, S., Makedonska, N., Gable, C. W., Painter, S. L., and Viswanathan, H. S. (2015), dfnWorks: A discrete fracture network framework for modeling subsurface flow and transport, *Computers & Geosciences*, 84, 10–19.

Islam, A. W., Patzek, T. W., and Sun, A. Y. (2015), Thermodynamics phase changes of nanopore fluids, *Journal of Natural Gas Science and Engineering*, 25, 134-139.

Liu, X., and Zhang, D. (2019), A review of phase behavior simulation of hydrocarbons in confined space: Implications for shale oil and shale gas, *Journal of Natural Gas Science and Engineering*, 68, 102901.

Lougheed, D., and Anderson, D. (2017), Does Pressure Matter? A Statistical Study. In SPE's *Unconventional Resources Technology Conference (UreTEC)*, Austin, Texas (pp. 1919-1933).

Mirani, A., Marongiu-Porcu, M., Wang, H., and Enkababian, P. (2018), Production-pressure-drawdown management for fractured horizontal wells in shale-gas formations. *SPE Reservoir Evaluation & Engineering*, 21(03), 550-565.

Mohaghegh, S. D., (2017), Shale Analytics: Data-driven Analytics in Unconventional Reservoirs, Springer, Switzerland.

MSEEL: http://mseel.org/

Muggeridge, A., Cockin, A., Webb, K., Frampton, H., Collins, I., Moulds, T., and Salino, P. (2014), Recovery rates, enhanced oil recovery and technological limits. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2006), 20120320.

Pan, S. J., and Yang, Q. (2009), A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Patzek, T. W., Male, F., and Marder, M. (2013), Gas production in the Barnett Shale obeys a simple scaling theory. *Proceedings of the National Academy of Sciences*, 110(49), 19731-19736.

PetroMehras -- List of Petroleum Softwares and Simulators, https://www.petromehras.com/petroleum-software-directory/reservoir-simulation-software/list-of-petroleum-software-application

PFLOTRAN: https://www.pflotran.org/

Puyang, P., Taleghani, A. D., and Sarker, B. R., (2015), Multi-disciplinary data integration for inverse hydraulic fracturing analysis: A case study. *SPE Journal*, 2015.

Rajput, S., and Thakur, N. K., (2016), Geological Controls for Gas Hydrates and Unconventionals, Elsevier, Cambridge, Massachusetts, USA.

Rezaee, R. (2015), *Fundamentals of Shale Gas Reservoirs*, John Wiley & Sons, New jersey, 2015.

Salama, A., Amin, M. F. E., Kumar, K., and Sun, S. (2017), Flow and transport in tight and shale formations: A review. *Geofluids.*

Sandrea, I. & Sandrea, R. (2007), Global oil reserves - 1: Recovery factors leave vast target for EOR technologies. *Oil Gas Journal*.

Seales, M. B., Ertekin, T., and Wang, J. Y. (2017), Recovery efficiency in hydraulically fractured shale gas reservoirs. *Journal of Energy Resources Technology*, 139(4), 042901.

Sun, H. (2015), *Advanced Production Decline Analysis and Application*. Gulf Professional Publishing, Elsevier, Cambridge, Massachusetts, USA.

Tan, S. P., and Piri, M. (2015), Equation-of-state modeling of confined-fluid phase equilibria in nanopores, *Fluid Phase Equilibria*, 393, 48-63.

Weijermars, R. (2014), U. S. shale gas production outlook based on well roll-out rate scenarios, *Applied Energy*, 124, 283-297.

Wu, X., He, J., Zhang, H., & Ling, K. (2013), Tactics and Pitfalls in Production Decline Curve Analysis. *In SPE Production and Operations Symposium*.

Yamada, M., Chen, J., and Chang, Y. (2018), *Transfer Learning: Algorithms and Applications*. Morgan Kaufmann.

Zitha, P., Felder, R., Zornes, D., Brown, K. & Mohanty, K. (2008), Increasing Hydrocarbon Recovery Factors. *SPE Journal*.

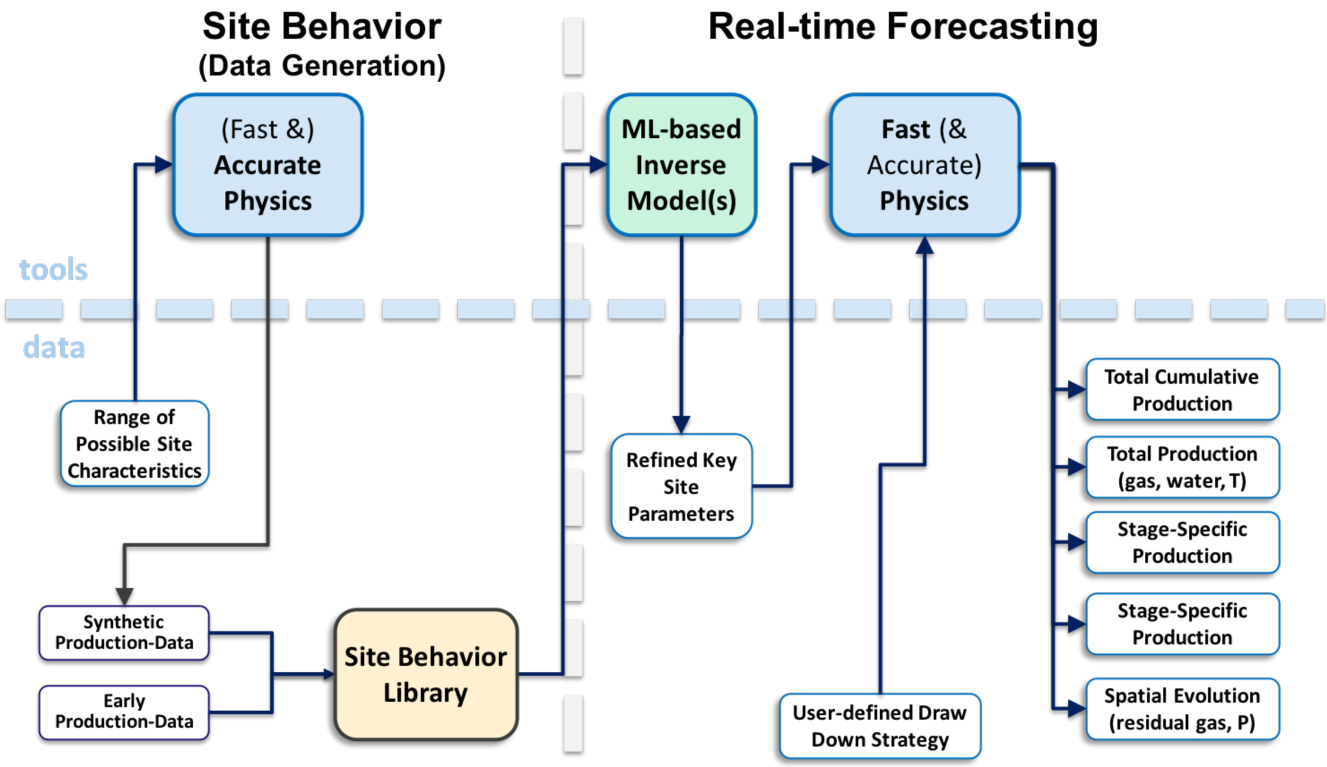

Figure-2: Details of PIML workflow. The workflow utilizes a library of data on site behavior to inform forecasting models. The scale of physical parameters used in the development of site behavior libraries are stage spacing (~100-200m), hydraulic fracture length (~100-150m), a stage may contain 3-4 hydraulic fractures and a swarm of natural fractures that are connected to hydraulic fractures, fracture rock reservoir permeability (~0.1-0.9mD), fractured rock porosity (~0.04-0.08), reservoir pressures (~20-30MPa), well flowing pressures (~5-15MPa), and gas properties (e.g., density, saturation, viscosity, compressibility).

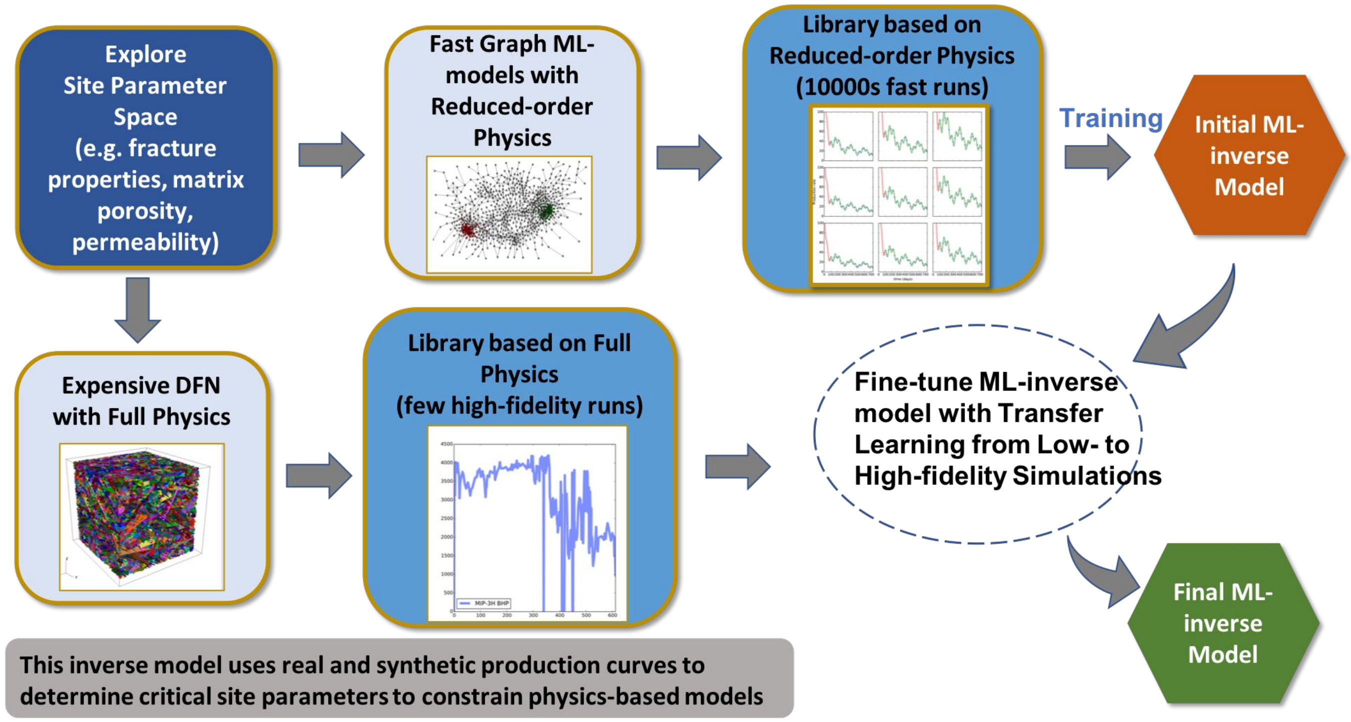

Figure-3: PIML workflow to create an efficient set of ML-inverse models for real-time history-matching.

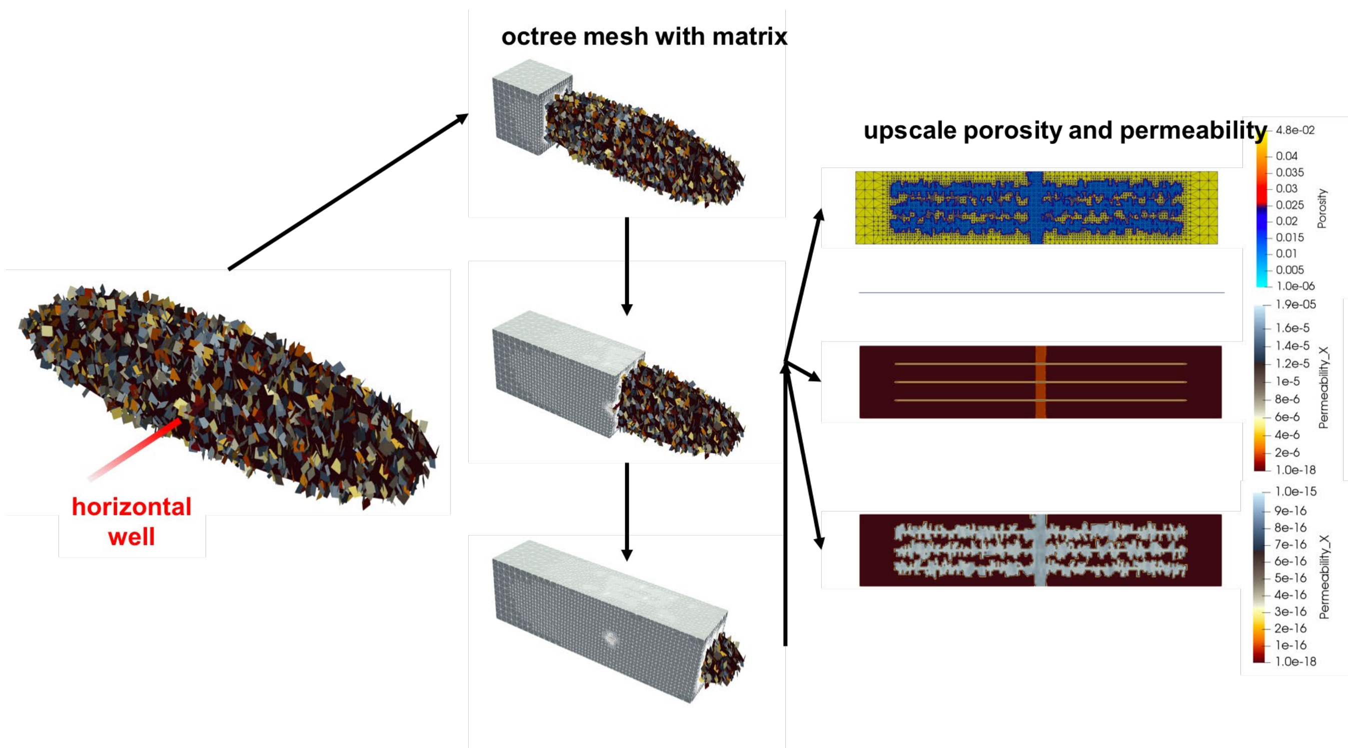

Figure-4: Discrete fracture network (DFN) and upscaled DFN models to simulate the physical behavior of the fractured reservoir.

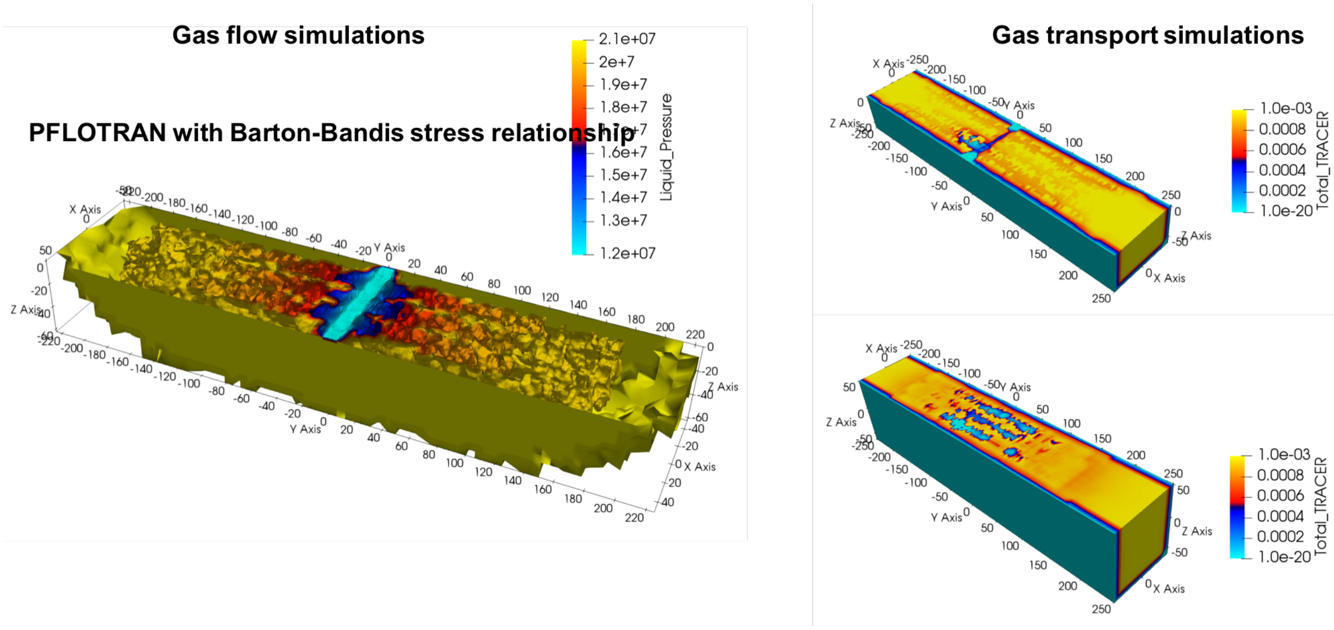

Figure-5: Gas flow and transport in upscaled DFN model using PFLOTRAN simulator with Barton-Bandis stress relationship.

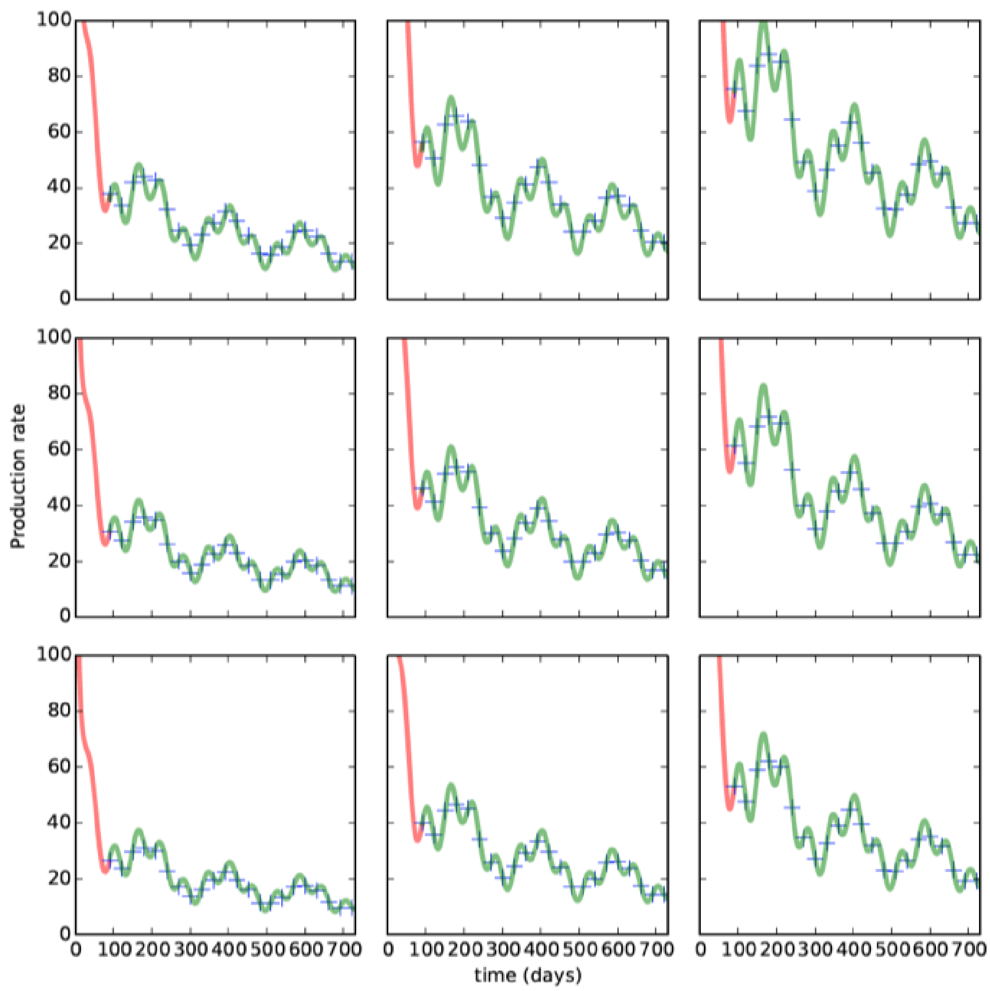

Figure-6: Long-term production forecasts using ML-forward model trained on reduced-order physics models.