DEVELOPING A MACHINE LEARNING FRAMEWORK TO REDUCE THE COMPUTATIONAL BURDEN OF REACTIVE TRANSPORT SIMULATIONS

Project Demographics:

Principal Investigator: Scott C. James
Project Title: "Developing a Machine Learning Framework to Reduce the Computational Burden of Reactive-Transport Simulations"
Project Period: 06/01/2019 to 08/31/2019
Department/School: Department of Geosciences
Email: sc_james@baylor.edu
Campus Phone: 254-710-2534
Amount Requested: $5,000.0

Non-technical abstract

Reactive transport modeling is required to make informed decisions about CO_2 sequestration, nuclear, and industrial waste disposal in the subsurface, etc. Geochemical reactions in multi-component reactive-transport modeling tremendously increase execution times, but the burdens can be reduced by including only the dominant geochemical reactions and ensuring that the primary characteristics governing mixing are included. Applying unsupervised machine learning (ML) tools to inputs and outputs of a reactive transport modeling can identify and extract only the dominant geochemical species to reduce execution times. Matrix factorization can be used to discover hidden features (in this case dominant geochemical species) in a database, but reactive-transport simulation requires multiple inputs and yields multiple outputs that result in a multidimensional database necessitating tensor factorization. Non-negative tensor factorization combined with customized k-means clustering, called NTFk, shows promise to extract dominant features from a multidimensional database. Moreover, NTFk is a structure-preserving feature-extraction method that honors the underlying physics of the chemical reactions. NTFk will be used to rank dominant geochemical reactions, understand the underlying mixing process related to each reaction, and quantify uncertainties of reactive-transport simulations. The outcome of this research will significantly reduce the computational burden of reactive-transport simulations for practical applications related to CO_2 injection, nuclear, municipal, and industrial waste storage, and deep-borehole brine injection.

1 Background and Rationale

Reactive-transport simulations are essential for making informed decisions about CO_2 injection, nuclear, municipal, and industrial waste storage, and deep-borehole brine injection (Bethke 2007; Lichtner, Steefel, and Oelkers 2019; Chen et al. 2018; Maes and Geiger 2018). Multi-component reactive-transport simulations are performed to estimate species decay rates, degrees of mixing, the extent of spreading, and product yields. However, execution times of transport models increase tremendously when geochemical species are considered but can be reduced by excluding higher-order geochemical species. Although discovering signatures of dominant chemical species is a complicated task, but it can be streamlined through machine learning (ML).

State-of-the-art ML methods use matrix (2D data) factorization to discover hidden signatures (Lee and Seung 1999; Cichocki et al. 2009; Vesselinov, Alexandrov, and O'Malley 2018; Iliev et al. 2018). Lee and Seung (1999) developed an algorithm, which learns parts of an object based on non-negative matrix factorization (NMF). The NMF algorithm is a contrast to principal component analysis and vector quantification that learn holistic not part based. NMF method has been applied successfully to a wide range of geological applications including discovering groundwater species, delayed signal, etc. (Iliev et al. 2018; Vesselinov, Alexandrov, and O'Malley 2018). But, reactive-transport signature discovery requires multiple inputs and yields multiple outputs, which result in a multidimensional signature-extraction database that is not amenable to analysis by matrix factorization (Alexandrov, Vesselinov, and Djidjev 2018). Tensor factorization is the only way to extract signatures (in this case geochemical species) from a multidimensional database. Vesselinov et al. (2018) used non-negative tensor factorization (NTF) with customized k-means clustering, NTFk, to study a simple chemical reaction to discover major geochemical signature and to better understand their underlying processes but not on a reservoir scale. This work will extend the work of Vesselinov et al. (2018) for a reservoir-scale model.

NTFk is a feature-extraction method from big data that is used in data analytics and data compression but sensitive to noise, so signature extraction strongly depends upon the initial guesses. Moreover, these methods require prior information by way of the number of signals to extract, which is a major drawback. The NTFk method overcomes these problems by extracting signatures subject to multiple initial guesses and then using a customized clustering procedure and sparsity constraints to identify the correct number of hidden signatures. NTFk provides robust estimates of the discovered signals even under noisy conditions, which makes it ideal for identifying dominant geochemical signatures under subsurface uncertainties. NTFk may also go through the non-unique optimization problem, which can be addressed through multi-start, regularizations, and non-negativity constraints (Alexandrov, Vesselinov, and Djidjev 2018).

NTFk is also a structure-preserving feature-extraction ML method, which means that it preserves the underlying physics of the chemical reactions (Alexandrov, Vesselinov, and Djidjev 2018). For example, concentrations of chemical species are non-negative. Under anisotropy and heterogeneity, finite-element and finite-volume simulations can produce oscillatory (even negative) concentrations of chemical species (Nakshatrala, Mudunuru, and Valocchi 2013). This nonphysical solution can be avoided with a structure-preserving feature-extraction method, that ensures non-negative concentrations through a robust constrained-

optimization technique (Cichocki et al. 2009). Multiple inputs and outputs produce massive multidimensional datasets (on the order of Terabytes). NTFk method scale linearly on state-of-art high-performance computing (HPC) systems (Vesselinov et al. 2018), which will be useful for analyzing massive multidimensional databases. In addition to using HPC systems, graphical/tensor processing unit computing, accounting data sparsity, and non-negative tensor training can also be used to deal with a large dataset. Given such computational efficiency, the NTFk method is attractive for signature discovery under heterogeneities at reservoir scales.

2 Specific Objectives:

The objective of this research is to develop an ML framework that will reduce the computational burden, enhance the understanding of the underlying physical processes, and quantify the uncertainties of reactive-transport simulations.

3 Potential Significance:

The outcome of this research will significantly reduce the computational burden of reactive-transport simulations for practical applications related to decision making for CO_2 injection, nuclear waste sequestration, deep-borehole brine injection, etc. In addition, this work will contribute to three distinct areas of reactive-transport modeling: (1) decrease the number of chemical species to build reduced-order models and simplified closed-form mathematical expressions, which can predict system behavior for less computational cost; (2) provide a better understanding of the underlying processes by interrogating the effects of mixing processes; and (3) determine the subsurface uncertainties (species concentration) on geochemical-signature discovery.

4 Plan of Work:

Phase I: Model development

The study will be conducted on the Farnsworth hydrocarbon unit (FU), which covers about 115 km^2 in northern Texas in the western Anadarko basin (Figure 1) (Khan 2017; Gallagher 2014). The target reservoir for CO_2 injection and enhanced oil recovery in the FU is the Morrow B Sandstone with 9 m thickness. The Morrow B lies at a depth of about 2,300 to 2,450 m, dipping to the southeast, at a temperature of about 75°C and an initial reservoir pressure of about 15.2 MPa. The porosity of the Morrow B varies from 10-18% and the permeability varies from 1 to 250×10^{-16} m^2. Hydrocarbons in the Morrow B are trapped stratigraphically by updip porosity pinchouts (McKay and Noah 1996).

A 3D numerical model will be developed to simulate the flow and reactive-transport of CO_2 injection into the reservoir using the thickness of the Morrow B Sandstone and overlying Morrow Shale using 1,000 × 1,000 × 7 cells. The thickness, porosity, and permeability variations of the Morrow B and Morrow Shale over the model domain were determined from well logs and geophysical surveys compiled by (Ampomah, Balch, and Grigg 2015) that will be used to represent heterogeneity. In this grid, the five lowermost layers will be assigned

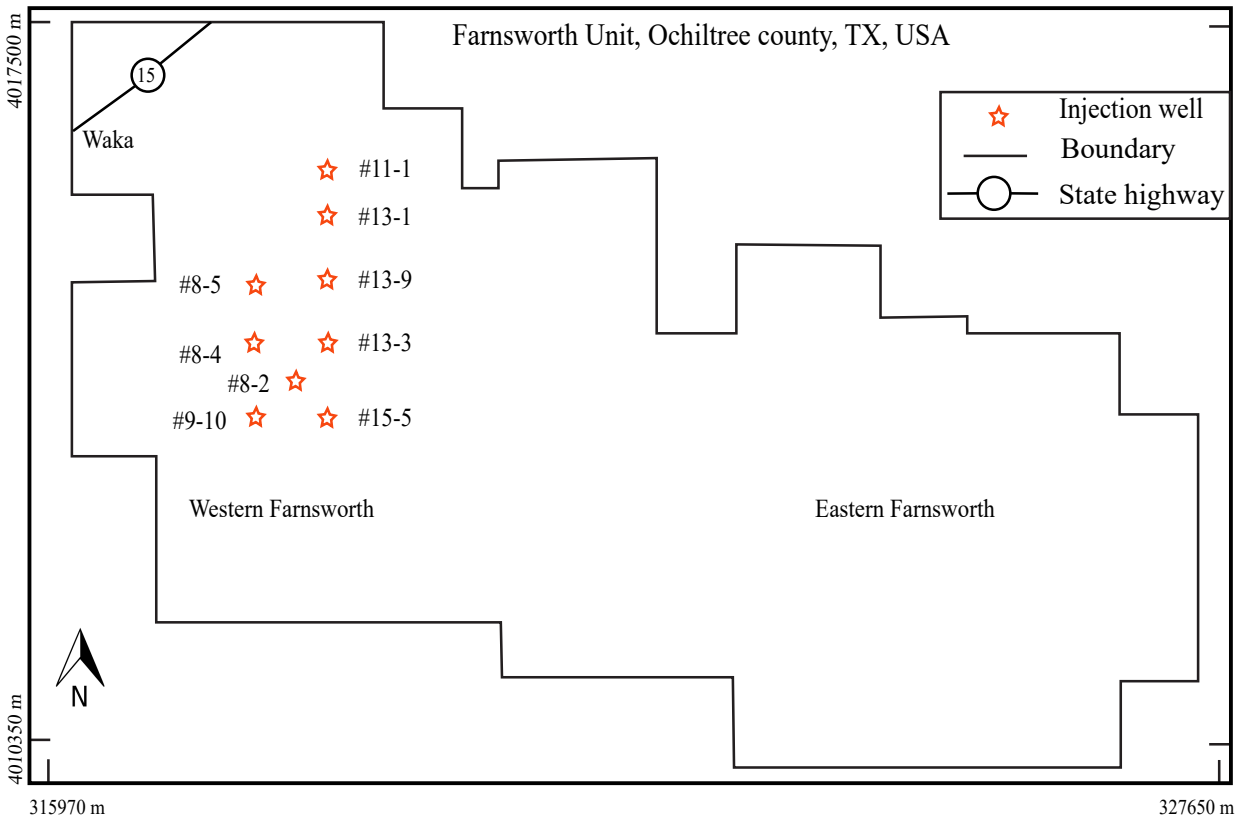

Figure 1: Plan view of the FU showing locations of CO_2 injection wells.

to the Morrow B Sandstone and the two uppermost layers will be assigned to the Morrow Shale.

Several further initial conditions will be prescribed. Fluid pressure is 15.2 MPa at the western boundary at the base of the Morrow B and it decreases eastward according to a regional hydraulic head gradient of 2.2×10^{-3} (assumed to be equal to the regional topographic gradient), and to decrease hydrostatically with increasing elevation. Thus, the resulting equation describing the initial fluid pressure distribution as a function of distance (x) eastward from the western boundary of the model domain and elevation (z) above sea level is:

$$P = \rho g \left(132 - 2.2 \times 10^{-3} x - z\right), \tag{1}$$

where P is fluid pressure (Pa), ρ is fluid density ($\mathrm{kg\,m^{-3}}$), and g is gravitational acceleration ($\mathrm{m\,s^{-2}}$). The model domain will assume to have an initially homogeneous temperature of 75°C and homogeneous formation water salinity of 3,600 $\mathrm{mg\,kg^{-1}}$. All porous media is fully saturated with water and a total of 11 minerals (Table 1) will be considered.

Several boundary conditions will be prescribed in the model. The top and bottom boundaries of the grid will be prescribed as no fluid flow boundary conditions open to solute transport. Fluid flow and solute will transport through lateral boundaries. Nine CO_2 injection wells shown in Figure 1 with field pumping schedules, real-screen depth intervals, and fluid will be assigned to the western part of the grid.

Table 1: Initial mineral composition of the Morrow B sandstone (Munson 1989; Gallagher 2014).

Mineral	Volume (%)	Mineral	Volume (%)	Mineral	Volume (%)
quartz	84.26	kaolinite	2.72	ankerite	0.25
anorthite	4.50	mont	0.10	calcite	0.75
albite	4.50	illite	0.88	magnesite	0.00
chlorite	1.78	siderite	0.25		

For geologic media with multiphase flow, effective permeability varies so the relative permeability (ratio between effective and absolute permeabilities) is a sensitive variable. Relative permeability will be defined according to the van Genuchten-Mualem model (Pruess, Oldenburg, and Moridis 1999) using a residual water saturation of 0.15, a residual CO_2 saturation of 0.1 (White et al. 2014), and the Mualam model parameter of $\gamma = 0.457$ (Pruess et al. 2001; Xu et al. 2006). The relationship between capillary pressure and saturation will be defined according to a van Genuchten function (Pruess, Oldenburg, and Moridis 1999) using a strength coefficient ($\alpha = 19.61\,\text{kPa}$) typical of sandstone (Pruess et al. 2001; Xu et al. 2006) and a maximum pressure of 15.2 MPa. Based on the recommendation of (Pruess et al. 2001), the residual water saturation will be set to zero for capillary pressure calculations to cancel its negative infinity in the Van Genuchten model as the relative permeability for water approaches zero when water saturation approaches residual water saturation.

The mutual solubilities of water and CO_2 in one another will be computed according to Spycher and Pruess (2005). (Four equations are necessary to represent mutual solubilities between water and CO_2 and entahlpies of brine and CO_2. Including all equations will make it too equation heavy because I am including few equations on the methods section. Let me know, if you still want these equations.) The dependence of water density on CO_2 concentration will be calculated according to Julio E Garcia (2001). The enthalpies of brine and CO_2 will be calculated according to Julio Enrique Garcia (2003). In the simulations, CO_2 will be injected at the reservoir temperature for 10 years at field rates, followed by a twenty-year post-injection period. Aqueous species and gaseous species diffusion coefficients will be set to 10^{-9} $m^2\,s^{-1}$ and 1.1×10^{-5} $m^2\,s^{-1}$, respectively. Pore compressibility will be set to 10^{-10} Pa^{-1}, an average value for sandstone (Ingebritsen and Appold 2012), and the tortuosity factor will be calculated to be 0.525 based on the 14.5% average porosity of the Morrow B.

This is tortuosity factor not tortuosity

Phase II: Machine Learning Framework Development

NTFk will be used to extract dominant geochemical species to decrease the number of chemical species to build reduced-order models for the less computational cost. Structure-preserving feature of NTFk will not affect the underlying process so it will allow a better understanding of the underlying processes by interrogating the effects of mixing processes. The proposed NTFk method can be split into four steps: Minimization, Elimination, Custom clustering, and Uncertainty Analysis.

Minimization

There are multiple tensor factorization methods (De Lathauwer, De Moor, and Vandewalle 2000) and among them, Tucker decomposition will be used for three-dimensional data, see Figure 2, that is generated by a series of geochemical reactions in time and space. Mathematically Tucker-3 decomposition of the 3D tensor $C(s,w,t)$:

$$C\left(s,w,t\right)=G\otimes W\left(s\right)\otimes H\left(w\right)\otimes V\left(t\right)+\epsilon\left(s,w,t\right) \tag{2}$$

where $\otimes$ denotes the tensor product. The decomposition of the non-negative tensor $C(s,w,t)$ ($C\in\mathbb{R}_{\geq 0}^{K\times M\times N}$) can be expressed by components:

$$C_{ijl}=\sum_{p=1}^{k}\sum_{q=1}^{m}\sum_{r=1}^{n}G_{pqr}W_{ip}H_{jq}V_{lr}+\epsilon_{ijl}\quad\forall i,j,l \tag{3}$$

where all the elements of C, G, W, H, and V are non-negative and (3) can be represented as:

$$C_{ijl}=\widetilde{C}_{ijl}+\epsilon_{ijl} \tag{4}$$

where $\widetilde{C}$ ($\widetilde{C}\in\mathbb{R}_{\geq 0}^{K\times M\times N}$) is the Tucker-3 estimate of C:

$$\widetilde{C}_{ijl}=\sum_{p=1}^{k}\sum_{q=1}^{m}\sum_{r=1}^{n}G_{pqr}W_{ip}H_{jq}V_{lr} \tag{5}$$

where s, w, and t represent geochemical species, well, and time, respectively; G ($G\in\mathbb{R}_{\geq 0}^{k\times m\times n}$) is an unknown core-tensor that represents the interactions between the s, w, and t components of $W(s),H(w)$, and $V(t)$; W ($W\in\mathbb{R}_{\geq 0}^{K\times k}$) is an unknown matrix representing the geochemical signature; H ($H\in\mathbb{R}_{\geq 0}^{M\times m}$) is an unknown matrix accounts for dependence on the monitoring points, and V ($V\in\mathbb{R}_{\geq 0}^{N\times n}$) is an unknown matrix representing captures time dependence; $\mathbb{R}_{\geq 0}$ denotes the set of non-negative real numbers $\mathbb{R}_{\geq 0}=\{x\in\mathbb{R}\mid x\geq 0\}$, i ranges from 1 to K where K is the number of geochemical species; j ranges from 1 to M where M represents monitoring well number; l ranges from 1 to N where N is the number of time frames (snapshots); $1\leq k<K$; $1\leq m<M$; and $1\leq n<N$. Additionally, ϵ ($\epsilon\in\mathbb{R}^{K\times M\times N}$) is the unknown error between the original data $C(s,w,t)$ and estimated $G\otimes W(s)\otimes H(w)\otimes V(t)$ that is caused by presence of a noise or errors in the measurements. For a successful Tucker decomposition, ϵ should represent a white noise and an important part of NTFk analysis is to make sure that ϵ characterizes a white noise and there additional signals that are not extracted by this procedure. NTF allows the tensor C to be sparse so C can have missing observations.

To extract unknown core tensor G, and factor matrices W, H, and V, NTF starts with a random initial guess for W, H, and V and proceeds by minimizing cost function, O, the Frobenius norm,

$$O=\frac{1}{2}||C-\widetilde{C}||_F^2 \tag{6}$$

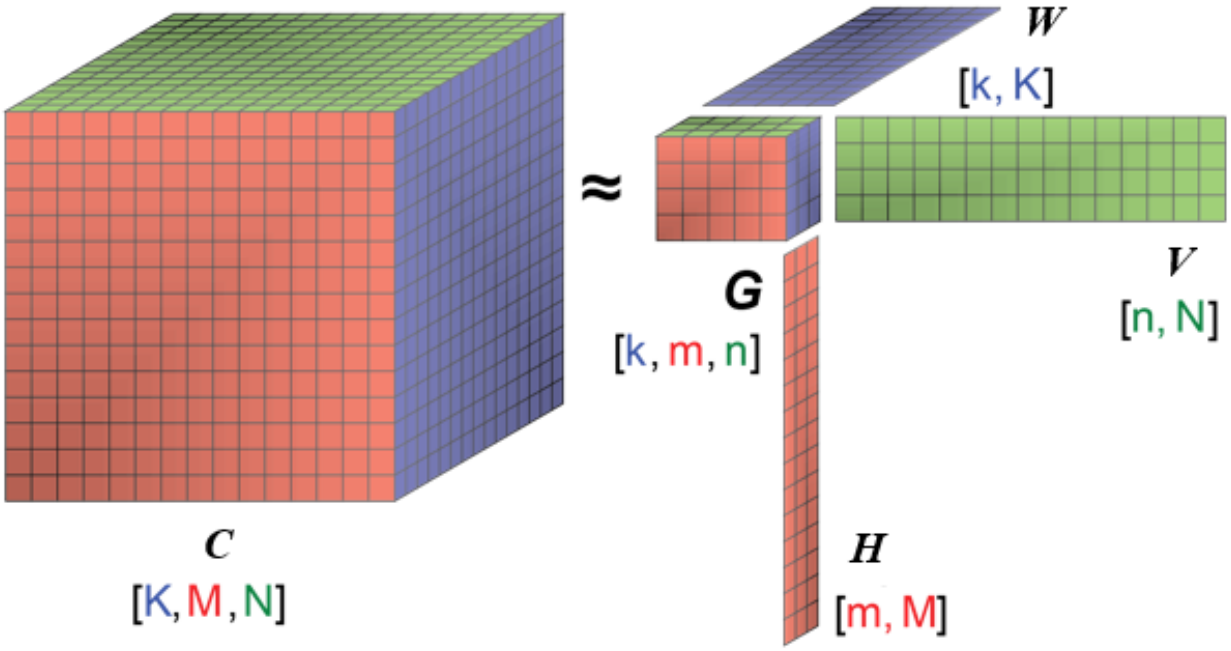

Figure 2: Schematic representation of Tucker based factorization for 3D tensors. Tucker-3 model decomposes the tensor $C(s,w,t)$ into a core tensor G and three factor matrices H, W, and V; modified from Vesselinov, Alexandrov, and O'Malley (2018).

during each iteration. The optimal number of features is evaluated by comparison of the quality of the reconstruction, O, and the quality of the derived clusters for different number of features, K, estimated by their average Silhouette value (Rousseeuw 1987), which will be discussed later.

Elimination criteria

The minimization depends on initial guesses of number of species, anisotropic-dispersion contrast, molecular diffusivity, spatiotemporal characteristics of the velocity field, and heterogeneities. The minimization may converge to a different solution or stop before getting a good solution, which makes it an ill-posed problem. To avoid unreasonable reconstruction of $\widetilde{C}$, 10% of worst reconstruction will be discarded.

Custom Clustering for determination of number of species

The number of original species, K, is required to explain the results of NTF method, but typically it is unknown. A custom clustering algorithm will be used to estimate the unknown number of species based on the robustness of the minimization solutions. The algorithm will consecutively examine all possible species of chemical reactions by obtaining a large number of NTF minimization solutions for each number of species. During clustering, the similarity

between two species will be measured using the cosine distance, as:

$$\rho(a,b) = 1 - \frac{\sum_{i=1}^{n} a_i b_i}{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2} \tag{7}$$

where a_i and b_i are individual components of the vectors a and b. After clustering, Silhoutte value (Rousseeuw 1987) will be calculated and use it to estimate a particular choice of K. Silhoutte value quantifies how similar an object is to its own cluster compared to other clusters and varies from -1 to +1, where high value indicates that the object is well matched to its own cluster and poorly matched to its neighbouring clusters. The combination of minimization and Silhoutte value will be used to determine the number of geochemical signatures. If K is low, Silhoutte value will be high but error of minimization will be high because solution would be under-fitted but for high K, Silhoutte value will be low and solution would be over-fitted. So, best estimate of K will be a number, which will optimize both minimization error and Silhoutte value.

Uncertainty analysis

We will use Bayesian analysis based on Markov Chain Monte Carlo (MCMC) sampling to obtain posterior probability distribution function (PDF). The PDFs will be obtained based on some species and their concentration using a likelihood function defined as $e^{-X^2/2}$. Bayesian analysis will provide a number of chemical species where the concentration of the original species can be found with a given probability.

Implementation

PFLOTRAN reactive-transport simulations will be run using 1,000 random initial guesses (ranging from worst to best) for model inputs (anisotropic-dispersion contrast, molecular diffusivity, and spatiotemporal characteristics of the velocity field) will be used to simulate $C(w,s,t)$ for each chemical species. This stacked tensor databases for model inputs and outputs will be input to the NTFk method, as described above, to identify hidden features (in this case model inputs for ADRE) and to quantify the contributions of these variable model inputs to product yield, species decay, degree of mixing, and spreading of chemical species across the domain. The NTFk analysis will be built using Python packages including Scikit-learn, Keras, Numpy, and Pandas.

Project Timeline:

Task	**Summer 2019**
Collect data for the FU and unsupervised ML approach	Weeks 1 and 2
Build the PFLOTRAN model to simulate reactive transport	Weeks 3 and 4
Build the ML framework for initial guesses and model heterogeneity	Weeks 5-7
Build the ML for framework for the reduced-order model	Weeks 8 and 9
Uncertainty quantification	Week 10

5 Specific Role of Undergraduate Researcher:

Undergraduate student researcher, First Last, will be responsible for the following:

- Collecting geologic (Pressure gradient, structure, location of injection and extraction wells), hydrogeologic (water pressure, porosity, and permeability, CO_2 injection overtime), and geochemical (species in the Farnsworth water) data
- Creating model grids with Farnsworth model dimension
- Overseeing model run for multiple initial guesses of model inputs
- Stacking inputs and outputs into tensor form and use them as input in NTFk algorithm.
- Computing uncertainties of model outputs
- Preparing a poster for the Hydrology Session at the American Geophysical Union Fall Meeting

6 Plans for Publication/Dissemination of Project Results:

All work on this project will be performed under the direct supervision of Prof. Scott James who is conducting research on Machine Learning (supervised and unsupervised). A manuscript for publication in the *Computational Geosciences* will be developed to include results from this research. Moreover, this work will be used to develop a poster for presentation at the American Geophysical Union Fall Meeting in the Hydrology Session.

References

Alexandrov, B., V. V. Vesselinov, and H. N. Djidjev. 2018. *Non-negative Tensor Factorization for Robust Exploratory Big-Data.* Technical report. LA-UR-18-20307, Los Alamos National Lab, Los Alamos, NM (United States).

Ampomah, W, RS Balch, and RB Grigg. 2015. "Analysis of upscaling algorithms in heterogeneous reservoirs with different recovery processes." In *SPE Production and Operations Symposium: SPE-173588-MS.* Society of Petroleum Engineers. https://doi.org/10.2118/173588-MS.

Bethke, C. M. 2007. *Geochemical and Biogeochemical Reaction Modeling.* Cambridge University Press.

Chen, Li, Mengyi Wang, Qinjun Kang, and Wenquan Tao. 2018. "Pore scale study of multiphase multicomponent reactive transport during CO_2 dissolution trapping." *Advances in Water Resources* 116:208–218.

Cichocki, A., R. Zdunek, A. H. Phan, and S.-I. Amari. 2009. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation.* John Wiley & Sons.

De Lathauwer, Lieven, Bart De Moor, and Joos Vandewalle. 2000. "A multilinear singular value decomposition." *SIAM journal on Matrix Analysis and Applications* 21 (4): 1253–1278.

Gallagher, Sara R. 2014. "Depositional and diagenetic controls on reservoir heterogeneity: upper morrow sandstone, Farnsworth unit, Ochiltree county, Texas." PhD diss., Citeseer.

Garcia, Julio E. 2001. "Density of aqueous solutions of CO_2." *Lawrence Berkeley National Laboratory,* no. 49023: 11.

Garcia, Julio Enrique. 2003. "Fluid dynamics of carbon dioxide disposal into saline aquifers." PhD diss., University California Berkeley, Berkeley, CA.

Iliev, Filip L, Valentin G Stanev, Velimir V Vesselinov, and Boian S Alexandrov. 2018. "Nonnegative Matrix Factorization for identification of unknown number of sources emitting delayed signals." *PloS one* 13 (3): e0193974.

Ingebritsen, Steven E, and MS Appold. 2012. "The physical hydrogeology of ore deposits." *Economic Geology* 107 (4): 559–584.

Khan, Riaz Hossain. 2017. "Evaluation of the Geologic CO_2 Sequestration Potential of the Morrow B Sandstone in the Farnsworth, Texas Hydrocarbon Field using Reactive Transport Modeling." Master's thesis, University of Missouri–Columbia, Columbia, Missouri.

Lee, Daniel D., and H. Sebastian Seung. 1999. "Learning the parts of objects by non-negative matrix factorization." *Nature* 401, no. 6755 (October): 788–791. ISSN: 0028-0836. doi:10.1038/44565. http://dx.doi.org/10.1038/44565.

Lichtner, Peter C, Carl I Steefel, and Eric H Oelkers. 2019. *Reactive transport in porous media.* Vol. 34. Walter de Gruyter GmbH & Co KG.

Maes, Julien, and Sebastian Geiger. 2018. "Direct pore-scale reactive transport modelling of dynamic wettability changes induced by surface complexation." *Advances in Water Resources* 111:6–19.

McKay, RH, and JT Noah. 1996. "Integrated perspective of the depositional environment and reservoir geometry, characterization, and performance of the Upper Morrow Buckhaults Sandstone in the Farnsworth Unit, Ochiltree County, Texas." *CIRCULAR-OKLAHOMA GEOLOGICAL SURVEY* 98:101–114.

Munson, Tim W. 1989. "Clay mineralogy and its controls on production, Pennsylvanian Upper Morrow Sandstone, Farnsworth Field, Ochiltree County, Texas." *AAPG Bulletin (American Association of Petroleum Geologists);(USA)* 73 (12).

Nakshatrala, K. B., M. K. Mudunuru, and A. J. Valocchi. 2013. "A numerical framework for diffusion-controlled bimolecular-reactive systems to enforce maximum principles and the non-negative constraint." *Journal of Computational Physics* 253:278–307.

Pruess, Karsten, Curtis M Oldenburg, and GJ Moridis. 1999. *TOUGH2 User's Guide Version 2.* Lawrence Berkeley National Laboratory.

Pruess, Karsten, Tianfu Xu, John Apps, Julio Garcia, et al. 2001. "Numerical modeling of aquifer disposal of CO_2." In *SPE/EPA/DOE Exploration and Production Environmental Conference.* Society of Petroleum Engineers.

Rousseeuw, Peter J. 1987. "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis." *Journal of computational and applied mathematics* 20:53–65.

Spycher, Nicolas, and Karsten Pruess. 2005. "CO_2-H_2O Mixtures in the Geological Sequestration of CO_2. II. Partitioning in Chloride Brines at 12–100 °C and up to 600 bar." *Geochimica et Cosmochimica Acta* 69 (13): 3309–3320.

Vesselinov, V. V., M. K. Mudunuru, S. Karra, D. O. Malley, and B. S. Alexandrov. 2018. "Unsupervised Machine Learning Based on Non-Negative Tensor Factorization for Analyzing Reactive-Mixing." *Journal of Computational Physics (in review), arXiv preprint arXiv:1805.06454.*

Vesselinov, Velimir V, Boian S Alexandrov, and Daniel O'Malley. 2018. "Contaminant source identification using semi-supervised machine learning." *Journal of contaminant hydrology* 212:134–142.

White, Mark D, Brian J McPherson, Reid B Grigg, William Ampomah, and Martin S Appold. 2014. "Numerical simulation of carbon dioxide injection in the western section of the Farnsworth Unit." *Energy Procedia* 63:7891–7912.

Xu, Tianfu, Eric Sonnenthal, Nicolas Spycher, and Karsten Pruess. 2006. "TOUGHREACT—a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: applications to geothermal injectivity and CO_2 geological sequestration." *Computers & Geosciences* 32 (2): 145–165.