Intelligent Ranking Tool for Analogue Classification in the Oil and Gas Industry

Stanford University

Cedric Fraces Gasmi (cfraces@Stanford.edu)

CS-229, Machine Learning (Fall 2017) Stanford University

Stanford | SCHOOL OF EARTH, ENERGY & ENVIRONMENTAL SCIENCES

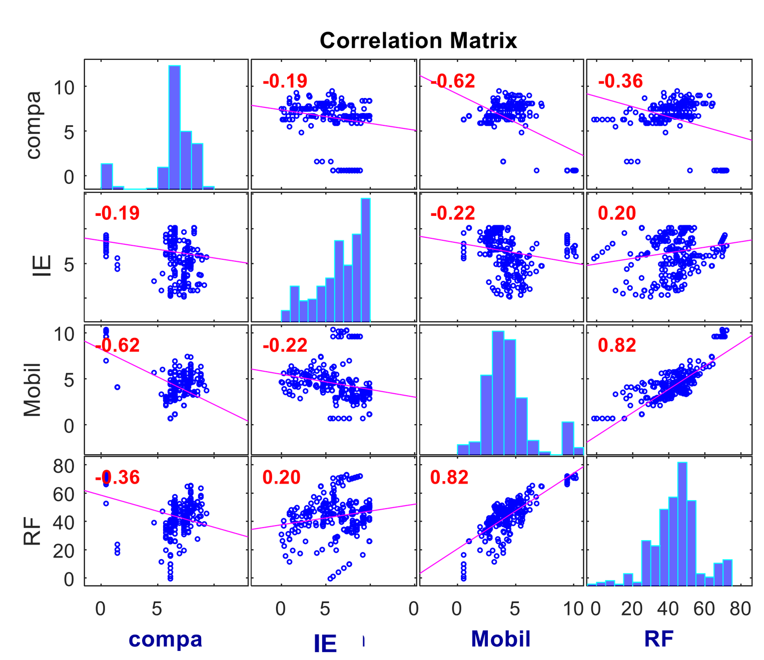
Background

The development of an oilfield is a lengthy industrial process that mobilizes large amount of capital investments and operator efforts in an uncertain environment. To mitigate such risks, experts very often resort to analogues for decision-making. An analogue is a reservoir that displays similar properties to the reservoir under study.

This project focuses on creating a tool for the systematic and robust referencing of reservoir analogues. Its purpose is to rank and evaluate reservoirs based on a set of key geologic parameters and give an estimate of their expected recovery (fraction of hydrocarbon in place that can be extracted from the reservoir). It is meant to be used by investors and decision makers who need to evaluate their portfolio and take rapid decisions on investment allocation. We explore different supervised and unsupervised learning techniques to compare the performances of data driven approaches with a standard correlation used in the industry. The results show a substantial improvement compared to classical methods.

Data

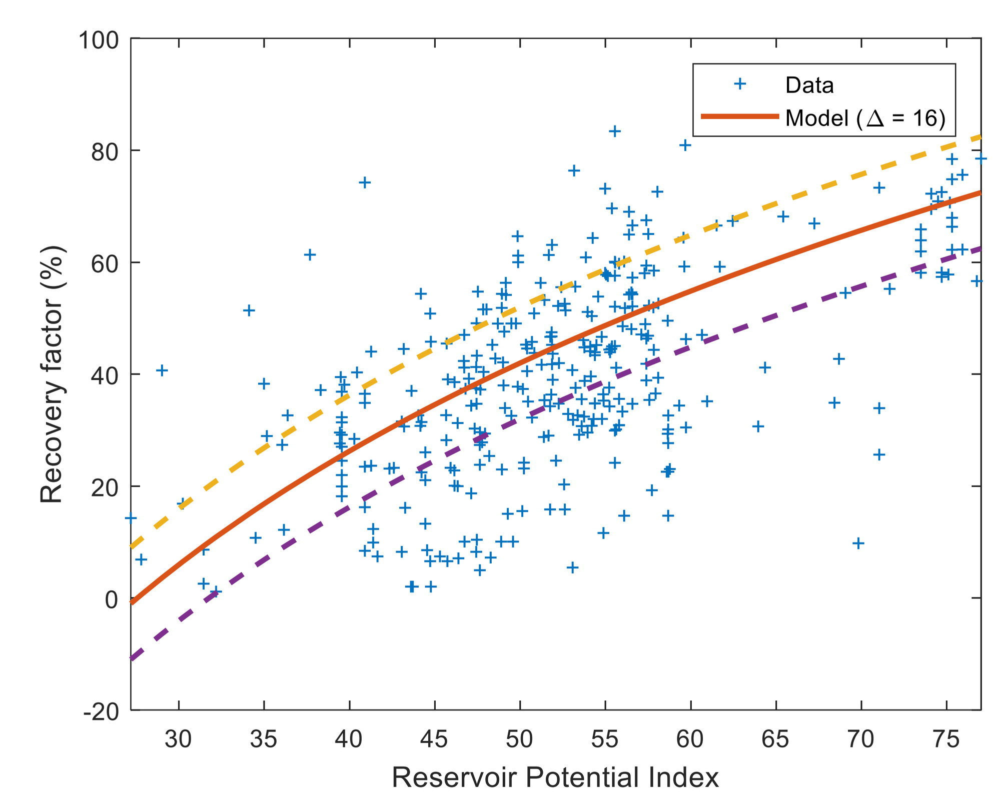
Public(*) and private databases containing geological features from various reservoirs with comprehensive sets of properties (porosity, permeability, viscosity, depth, fracture and fault level, depositional environment, Net to Gross ratio). These features are then grouped into "macro" categories: Compartmentalization, Intrinsic Energy, Mobility. These help lower the dimensionality of the problem and add some physics to the model



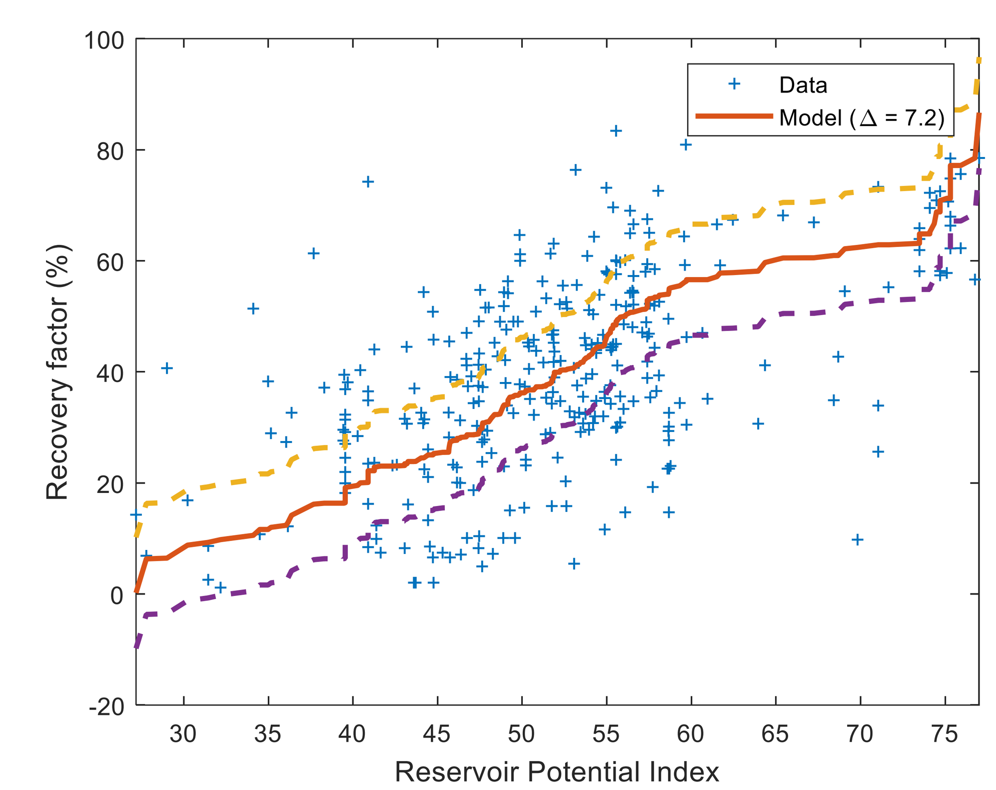
(*) http://www.rrc.state.tx.us

(*) http://www.conservation.ca.gov/dog

Standard Industry Approach



NNC



Methods

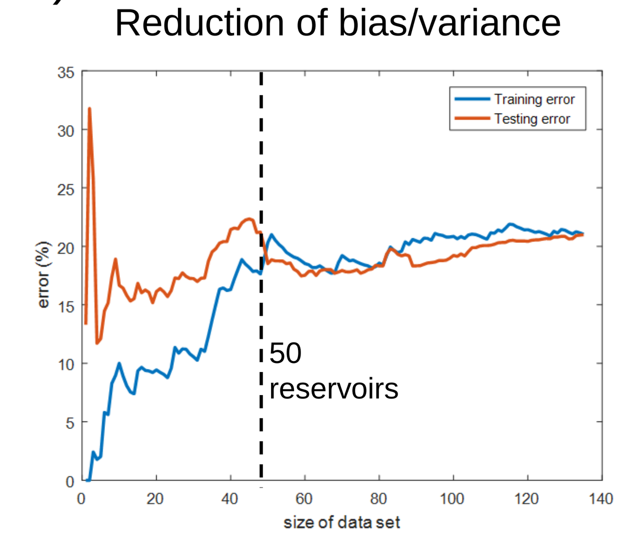
Unsupervised Learning

Clustering of reservoirs based on micro and macro geological features using k-means, NN and EM. Different combination of features and algorithms were tested. Due to the size of dataset and parameters to test, a model with 3 clusters and 3 macro features was chosen. It helped enhance the accuracy of the regression models.

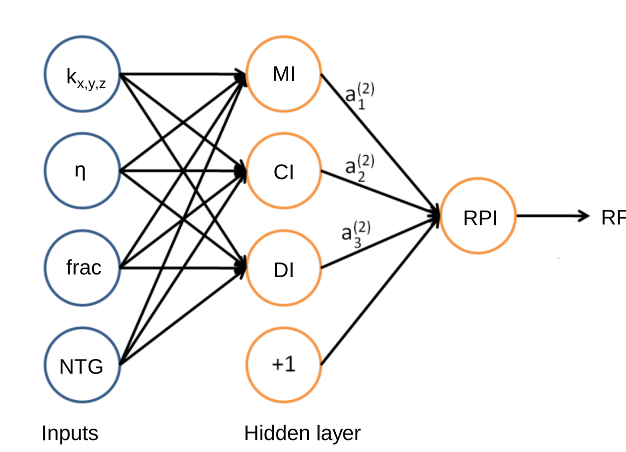
Supervised Learning

Linear and non linear regressions along with a 3 layers neural net were used to fit the model. The non linear regression was based on log series and trained on a subset of the whole data. Neural networks with various architectures (number of inputs and hidden layers) with Bayesian regression and Levenberg-Marquardt fit were tested. Training set (70%), Dev Set (15%), Test Set (15%)

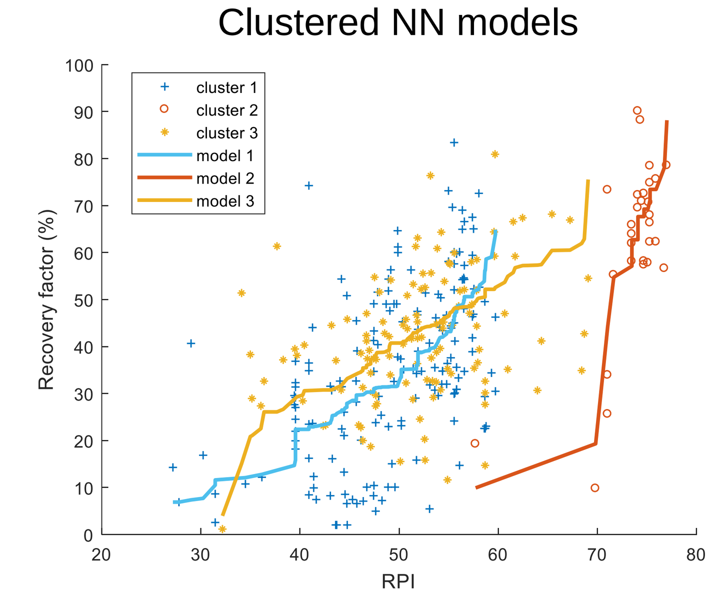
Reduction of bias/variance



Neural Net Architecture



Clustered NN models



Objective Function

We monitor the behavior of the model not only in a "one curve fits all" scenario but also when we operate within a range to account for technical and execution uncertainties.

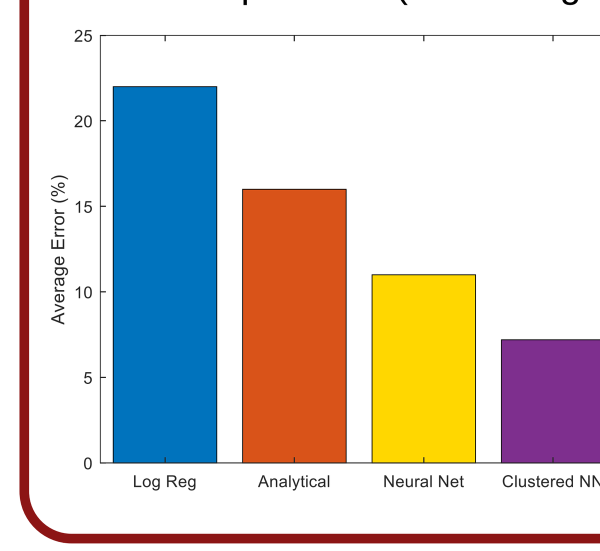
$$L(\theta) = \frac{1}{m}\sum_{i=1}^{m} \mu_i \underbrace{\left|h_\theta\left(x^{(i)}\right) - RF^{(i)}\right|}_{\Delta} \qquad \mu_i = \begin{cases} 0\ if\ \Delta \leq \overline{\delta} \\ 1 if\ \Delta > \overline{\delta} \end{cases}$$

margin

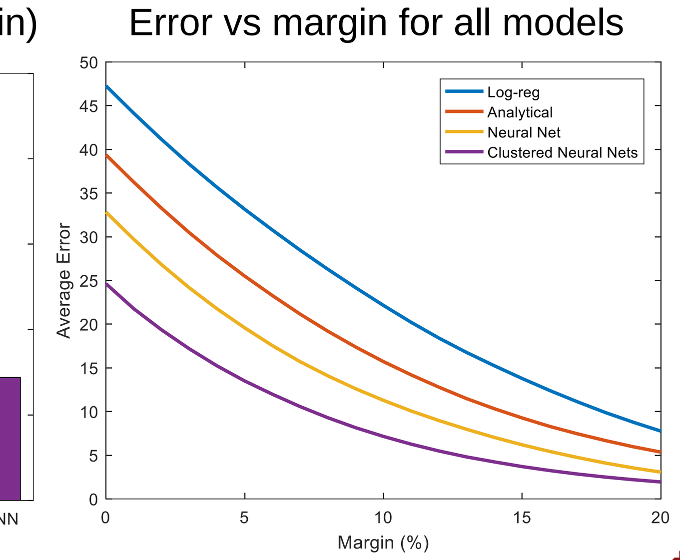
Results

We compute and compare the L1-errors for models with a +/-10% margin. We also run sensitivities w.r.t. the width of the margin. Neural nets alone show a 2 fold reduction in error (11% vs 22%). We then show additional improvement with the combination of Neural Networks within each cluster. Our combined model achieves a 3 fold reduction in error on average compared to a "basic" regression approach (7% vs 22%).

Error comparisons (10% margin)



Error vs margin for all models



Discussions

The study focused on the classification and ranking of analogue reservoirs purely based on geological features.

The method proposed achieves a drastic improvement in accuracy of prediction compared to industry standards (average error divided by 3) and naturally identifies the dependencies between features and macro-parameters.

Work to get additional data or augment the dataset using simulation should follow and will help test other hypothesis (unconventional, very mature fields).

The next step should focus on the realization of this potential recovery through the integration of the operational/financial components of the project. A similar approach based on the identification of "big hitters" can be extended. The coupling of these models will be a useful tool for a rapid screening of assets.

[1] Dake, L. (1978). Fundamentals of Reservoir Engineering. Amsterdam: Elsevier.

[2] G.S. Bromhal, J. B. (2014). Evaluation of rapid performance reservoir models for quantitative risk Assessment. Energy Procedia.

[3] H. Martin Rodriguez, E. E. (2013). New Approach to Identify Analogous Reservoirs. Society of Petroleum Engineers.

[4] Jaime Moreno, O. G. (2014). EOR Advisor System: A Comprehensive Approach to EOR Selection. International Petroleum Technology.

[5] Mohaghegh, S. D. (2016). Determining the main drivers in hydrocarbon production from shale using advanced data-driven analytics – A case study in Marcellus shale. Journal of Unconventional Oil and Gas Resources.