Automated Analysis of Temperature Data in Fractured Wells

Dante Isaac Orta Alemán (dorta@stanford.edu)[1]

[1]Department of Energy Resources Engineering, Stanford University

Stanford EARTH

PREDICTING

- The Oil & Gas industry has not yet fully understood the physics of fluid flow in hydraulically fractured wells. Understanding it is key to for operations.
- Temperature data can theoretically be used to estimate fluid flow from fractures.
- An evaluation of the potential of using Machine Learning for fluid flow estimation was done using synthetic temperature data. The results are promising and encourage future research.

MODELS & FEATURES

A) REGRESSION MODELS

I. Lasso Regression

$$\min_{\alpha}\left\{\sum_{i=1}^{N}\left(y_i-\sum_{k=1}^{K}\alpha_k T_k(x_i)\right)^2+\lambda\cdot J(\alpha)\right\},\quad J(\alpha)=\sum_{k=1}^{K}|\alpha_k|$$

II. Random Forest

$$\hat{f}_{\mathrm{rf}}^{B}(x)=\frac{1}{B}\sum_{b=1}^{B}T_b(x)\qquad T_b(x)=\sum_{m=1}^{M}c_m I(x\in R_m)$$

B) MODEL SELECTION

- The hyper parameters of each model were chosen through a 10-Fold Cross-Validation process
- **Lasso:** Regularization parameter
- **Forest:** Number of trees

C) FEATURE ENCODING

From the data variables set: {Temperature, Time, Length (x) } the following features were derived:

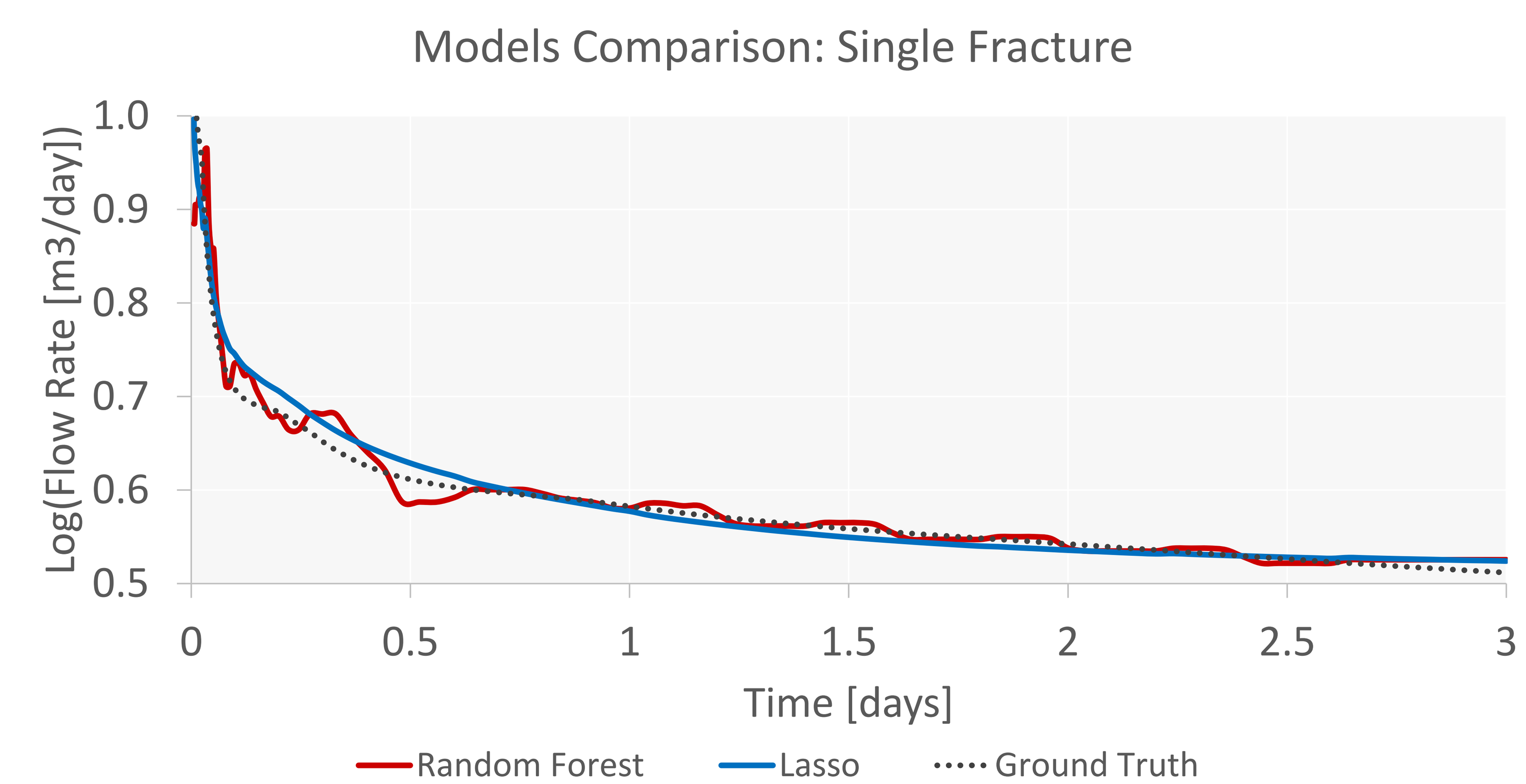
- Temperature
- Temperature spatial derivative
- Temperature temporal derivative

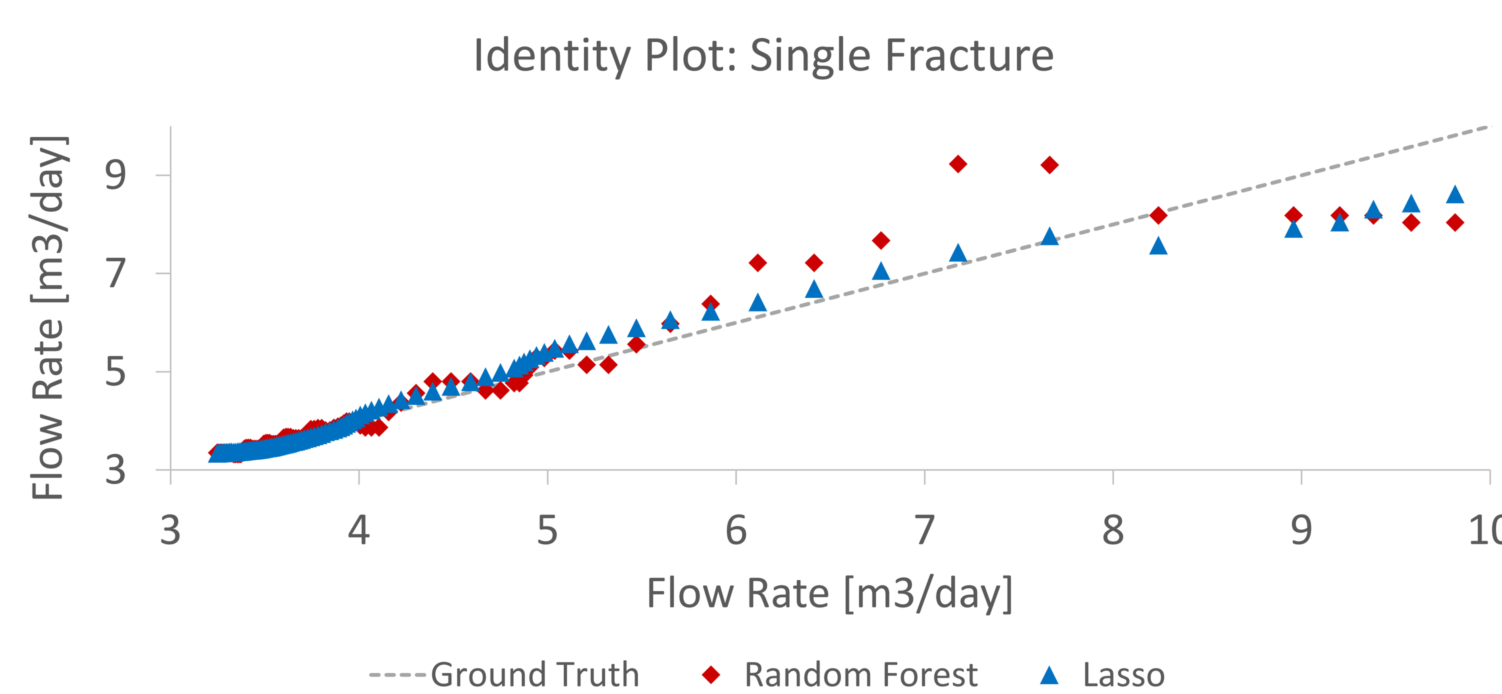
- A polynomial feature expansion was done up to the 3rd power
- Doing this proved to improve the model's performance
- For the lasso regression, the target variable was: log(Flow Rate)

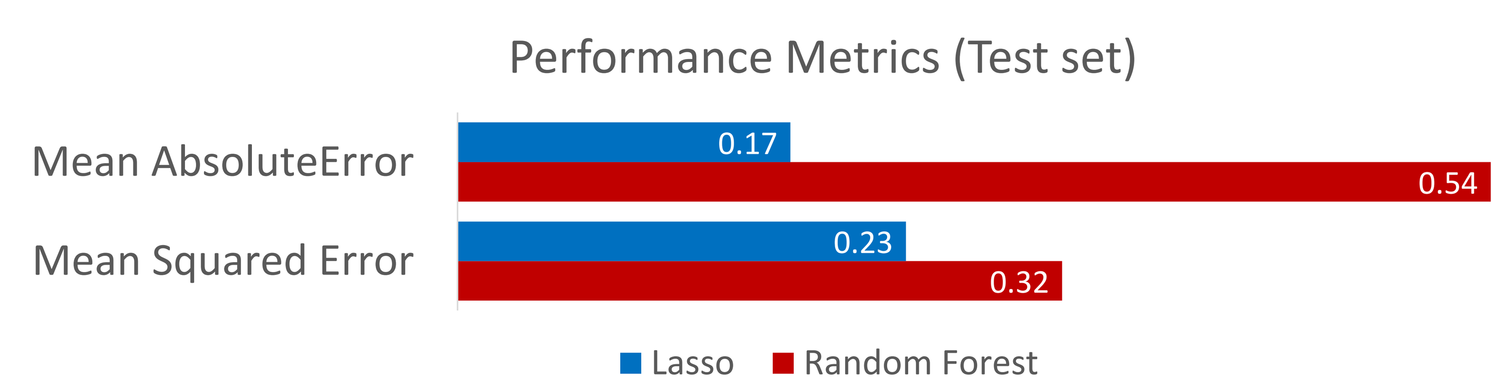
DATA

- The dataset comes from simulations of water injection processes at different flow rates into a multi-fractured well. The used simulator is the AD-GPRS flow simulator.
- The simulation outputs the following variables:
  - Spatial coordinates (x,y,z)
  - Spatial location of the fractures
  - Temperature at each location
  - Time step of the simulation
  - Flow rate at each fracture

RESULTS







DISCUSSION

- The Lasso regression proved to be the best model of the two tested.
- It was possible to model non-linear behavior using a linear model like the lasso by introducing a polynomial feature expansion .
- Both the Lasso and the Random Forest are capable of managing irrelevant features, thus allowing for increasing the dimensionality of the features.
  - The lasso achieves this through L1 regularization.
  - Each tree in the forest achieves this through entropy maximization.
- The Random Forest proved to have a noisier prediction. This is probably due to the nature of the piecewise constant functions that is composed by.
- A key step for improving the performance in the lasso, was to change the target variable to the logarithm of the flow rate.
- The random forest does not require any change in target variables to achieve peak performance. This is due to the invariance of trees to monotonic variable transformations

FUTURE WORK

- The models require testing in different injection/extraction conditions to test if the algorithm is capable of capturing the underlying physics.
- As the process is highly non-linear, there is the potential to improve the accuracy by changing the model. Good candidates for this are Neural Networks and Support Vector Machines

REFERENCES

1. Ribeiro, P. M., & Horne, R. N. (2016). Detecting Fracture Growth Out of Zone by Use of Temperature Analysis. Society of Petroleum Engineers. doi:10.2118/170746-PA
2. Hastie, T., Hastie, T., Tibshirani, R., & Friedman, J. H. (2001). The elements of statistical learning: Data mining, inference, and prediction. New York: Springer.