



Oh, Well.

A predictive analysis of rate
of penetration as a function
of drilling parameters of
offshore wells

Table of contents



01

A Solution for Chevron

Challenges facing Chevron

02

Progression of Analysis

How we approached the processing of our data

Our Models

A variety of models we explored and finding our favorite

03

Tuning and Testing

Final touches and testing our predictions

04

01

A Solution for Chevron

The impetus for our analysis



A Solution for Chevron



Offshore drilling requires investing in extremely specialized and costly resources.

Minimizing time can result in huge savings for the data-savvy.

02

Progression of Analysis

Our workflow



Progression of Analysis

0.

Data Cleaning

1.

Exploring Scope of
Data and Cartesian
Products

2.

Segmenting Data

Fig. 1. Data Exploration Formation ID

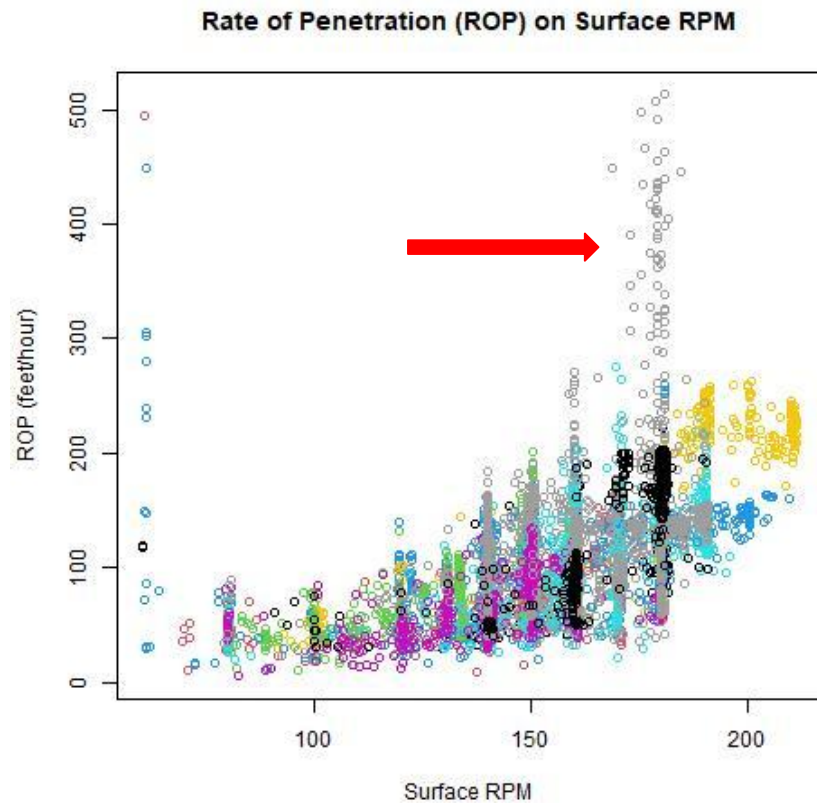
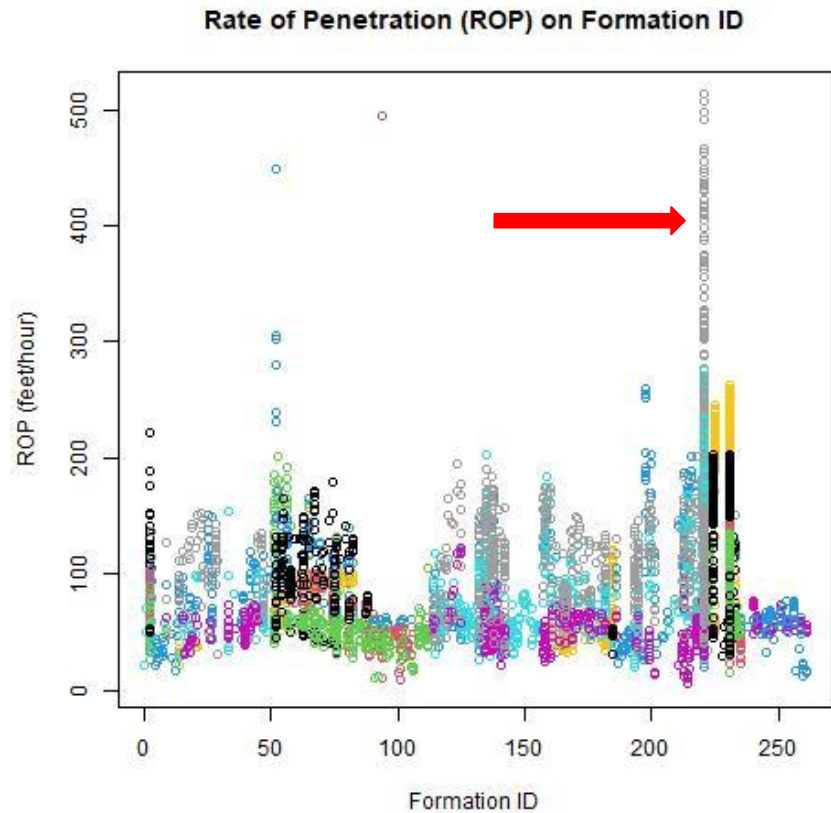


Fig. 2. Data Exploration Surface RPM

Data Exploration



03

Our Models

1.

Lasso

2.

Forward
stepwise
linear
regression

3.

Random
forest



Preliminary Factor Analysis

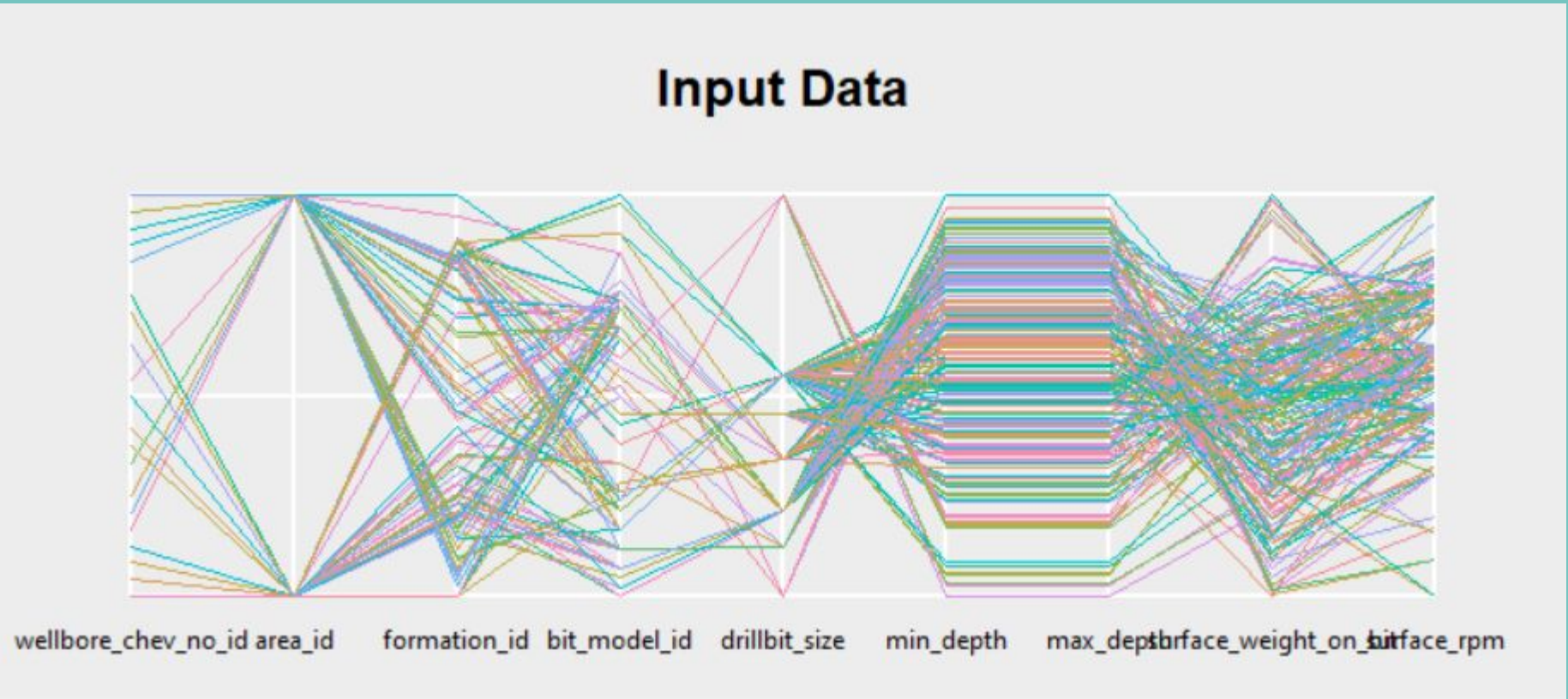


Fig. 3. Parallel Coordinate Plot of Potential Regressors

Preliminary Factor Analysis

Local Importance Scores

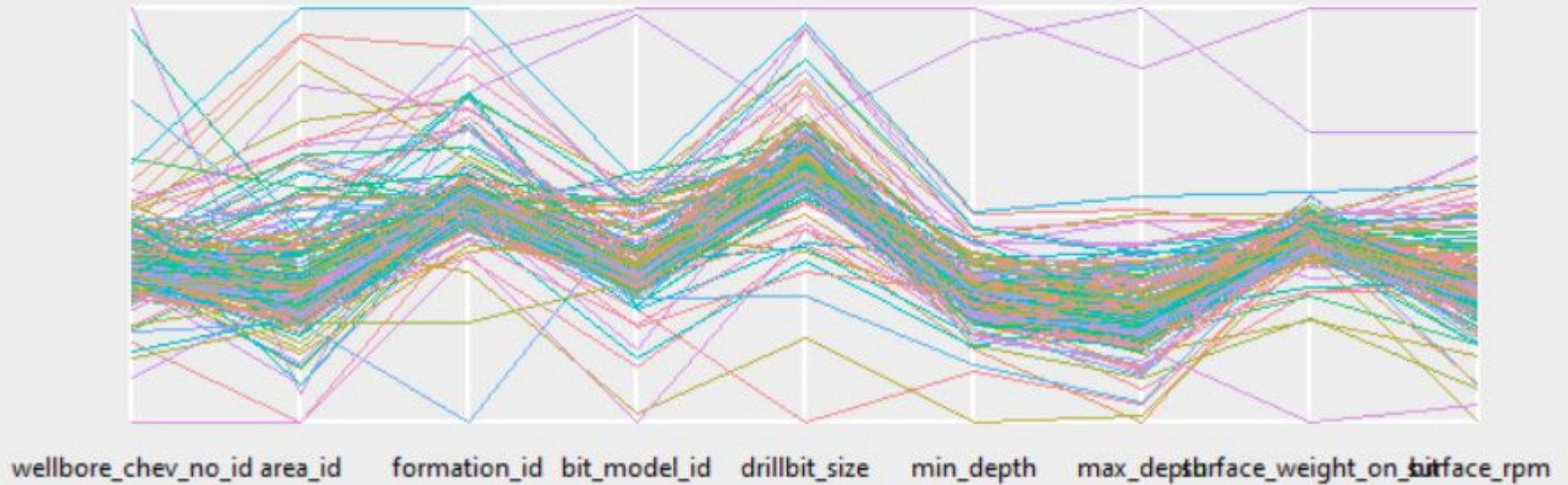


Fig. 4. Parallel Coordinate Plot of Local Importance

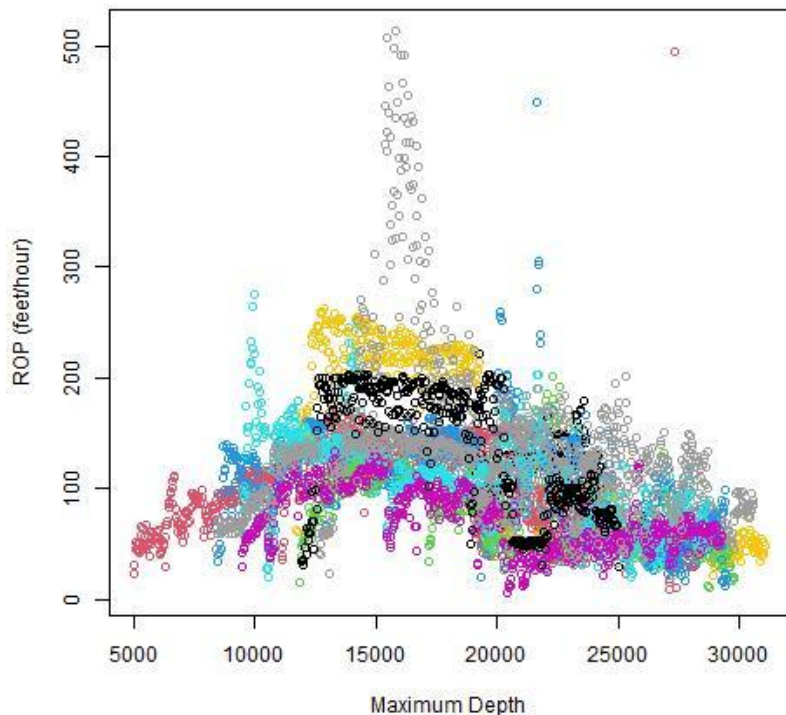
Forward stepwise linear regression

Model Iteration	Variable Selection	Variable Addition	R-Squared Adj.
1	Surface RPM	Surface RPM	0.3741
2	Surface RPM + Surface Weight on Bit	Surface Weight on Bit	0.4423
3	Surface RPM, Surface Weight on Bit	Minimum Depth	0.4595
4	Surface RPM, Surface Weight on Bit, Minimum Depth	Bit Model ID	0.4675

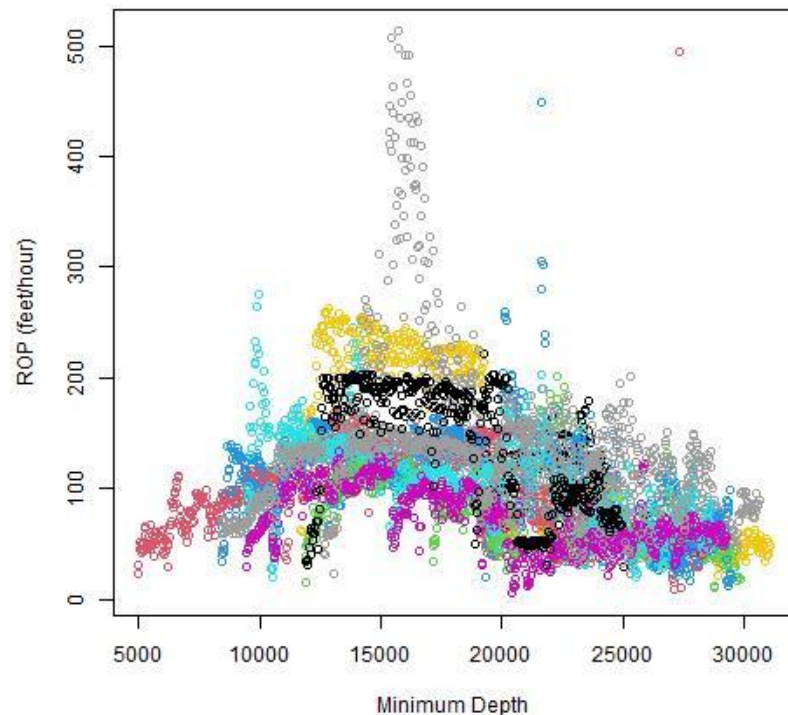
Due to concerns about multicollinearity between predictors, Maximum Depth was not considered for inclusion as a regressor.

Collinearity Analysis

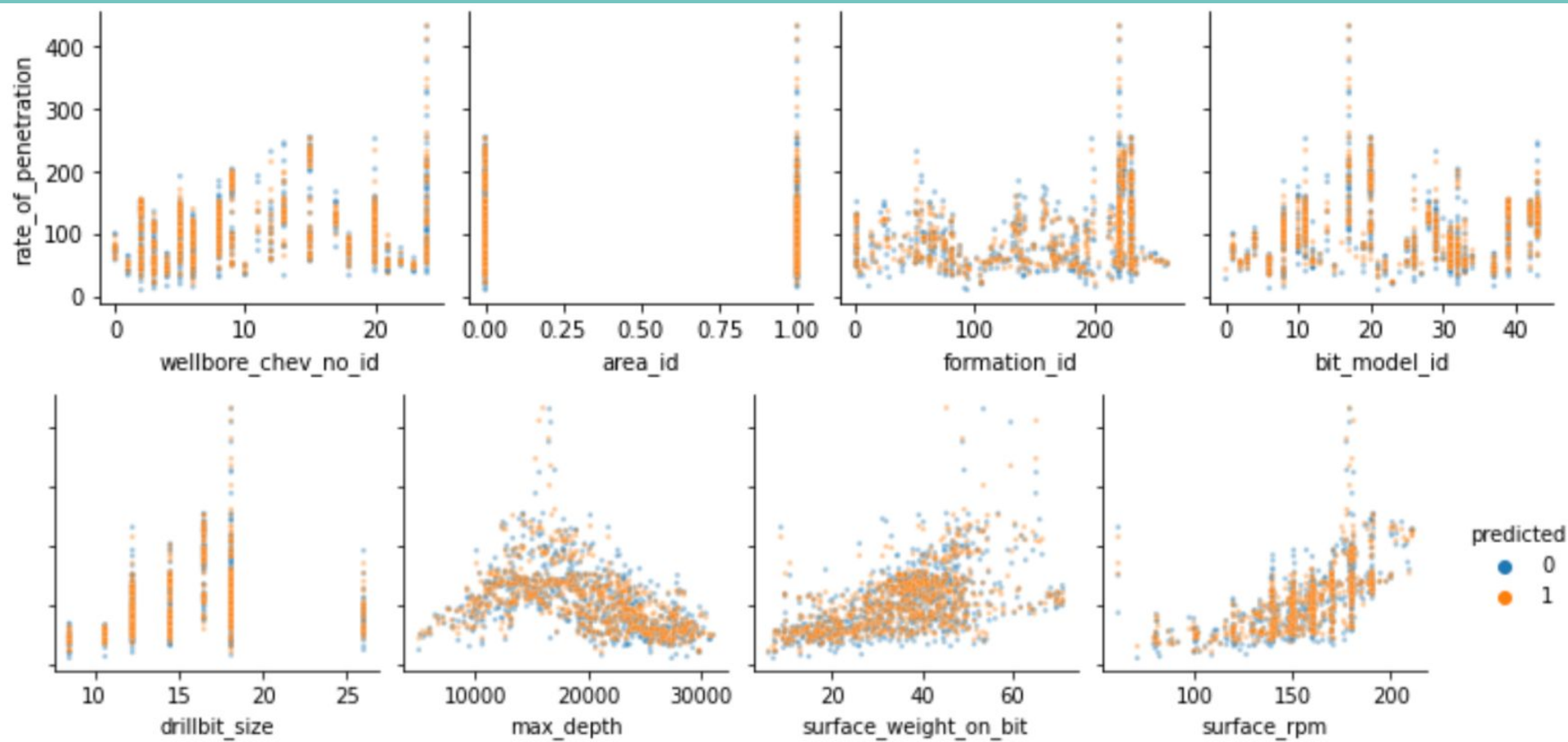
Rate of Penetration (ROP) on Maximum Depth



Rate of Penetration (ROP) on Minimum Depth



Performance of Random Forest Model predicting Rate of Penetration



04

Tuning and Testing

Hyperparameter optimization
and testing



Conclusions

- **Lasso Regression RMSE:** 29.6866
- **Forward Stepwise Linear Regression RMSE:** 40.16
- **Initial Random Forest Regression RMSE:** 16.1922
 - ◆ validation segment of the given data
 - ◆ without hyperparameter optimization via cross-validated grid-search

- **Final Random Forest RMSE:** 15.9985
 - ◆ *Test segment of given data*
 - ◆ with hyperparameter optimization via cross-validated grid-search