

FAILURE PREDICTION & EQUIPMENT RELIABILITY

Mohammad Evazi
Data Scientist, California Resources Corporation



Agenda

- About CRC
- Introduction to Data Science
- Applying machine learning to sucker rod pumps
- Dynamometer card classifier
- Well failure prediction and root-cause analysis
- Way forward

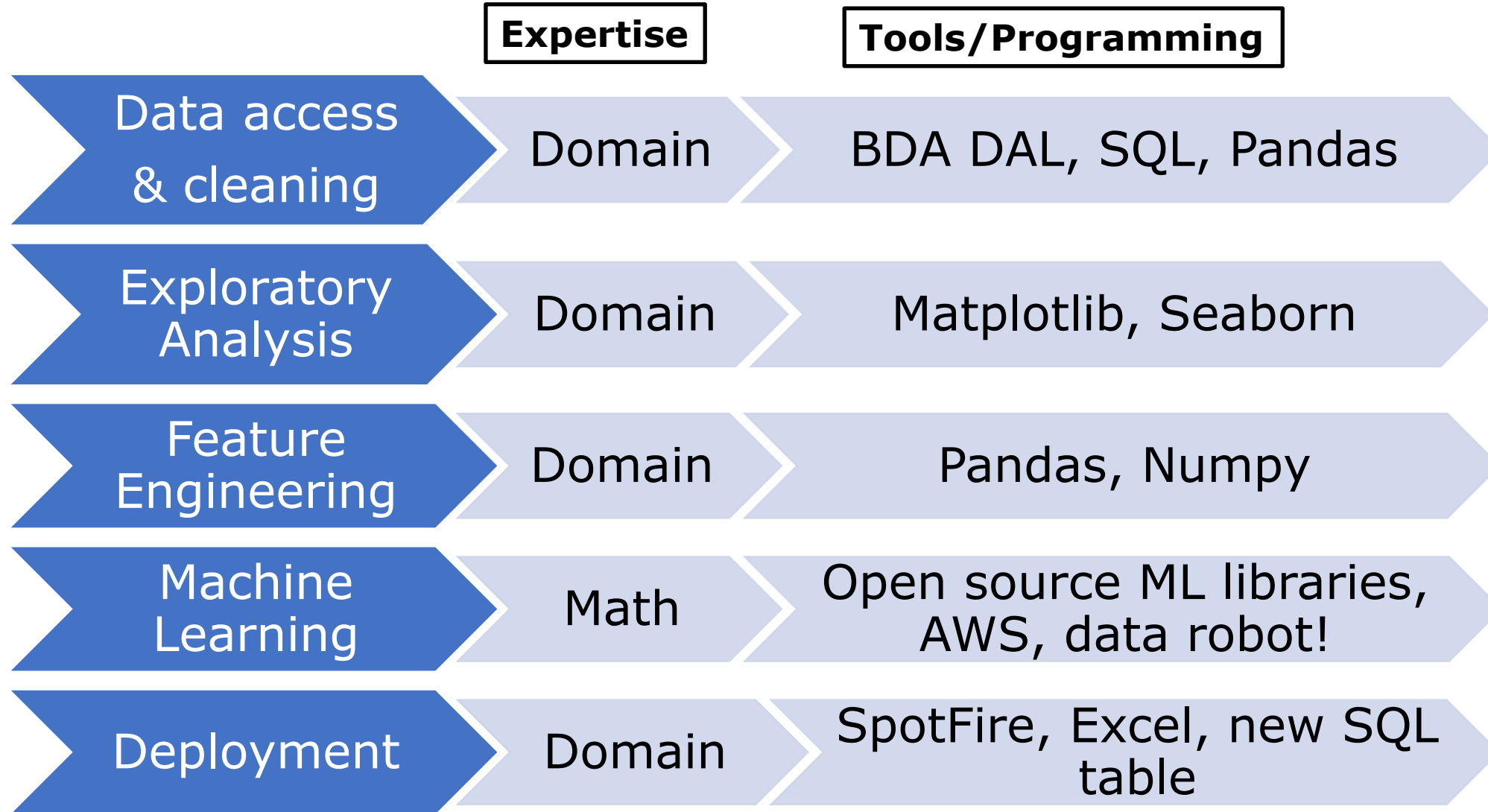
About CRC

- California Resources Corporation (CRC) is an oil and gas exploration and production company operating properties exclusively in the state of California
- 2014 spin-off of Occidental Petroleum
- Operates in San Joaquin, Los Angeles, Ventura and Sacramento basins
- Produces 134 MBOE/d with 60% oil
- Big Data Analytics Team formalized in late 2017

Data Science

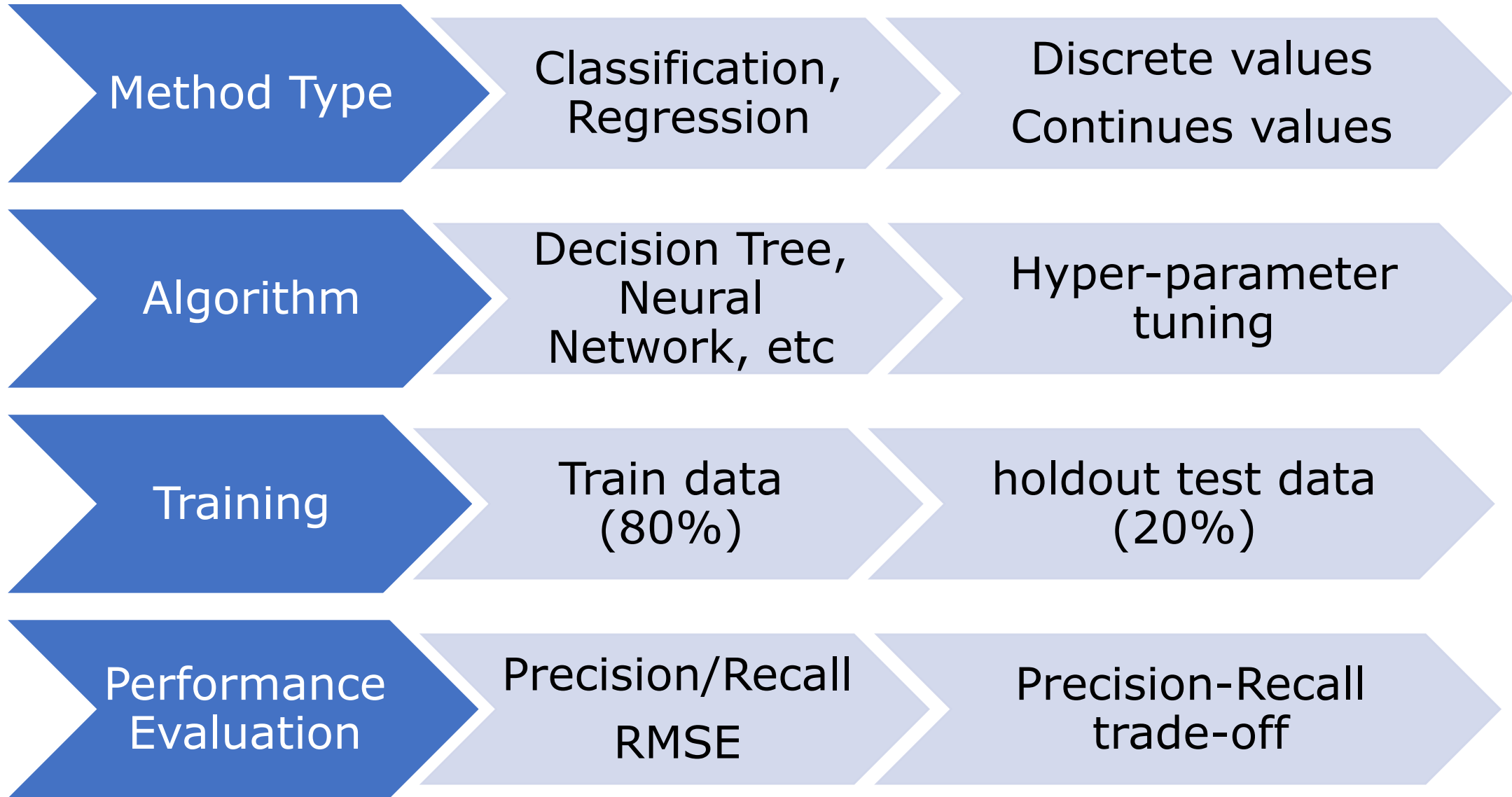
- No consensus on the definition!
- Mathematics + Programming + Domain Expertise (physics)
 - *These sets of tools provide a framework that can solve some of the data-intensive problems that we couldn't tackle before*
- Data science requires an innovation mindset; it takes time and iteration to succeed
- Early engagement of domain experts and agile development to establish quick-wins

Data Science Workflow - Tools



Programming: Python, Matlab, R

Building a Machine Learning Model



Well Failure Prediction: Size of the prize

- About 1,000,000 oil wells worldwide on sucker rod pump
- Annual failure rate of 0.2-0.6 per well
- Average failure cost is about \$30K
- Related downtime and oil production lost

Project Progression

Nov.
2017

Feb.
2018

April
2018

July
2018

Nov.
2018

BDA kick-off

BDA rollout

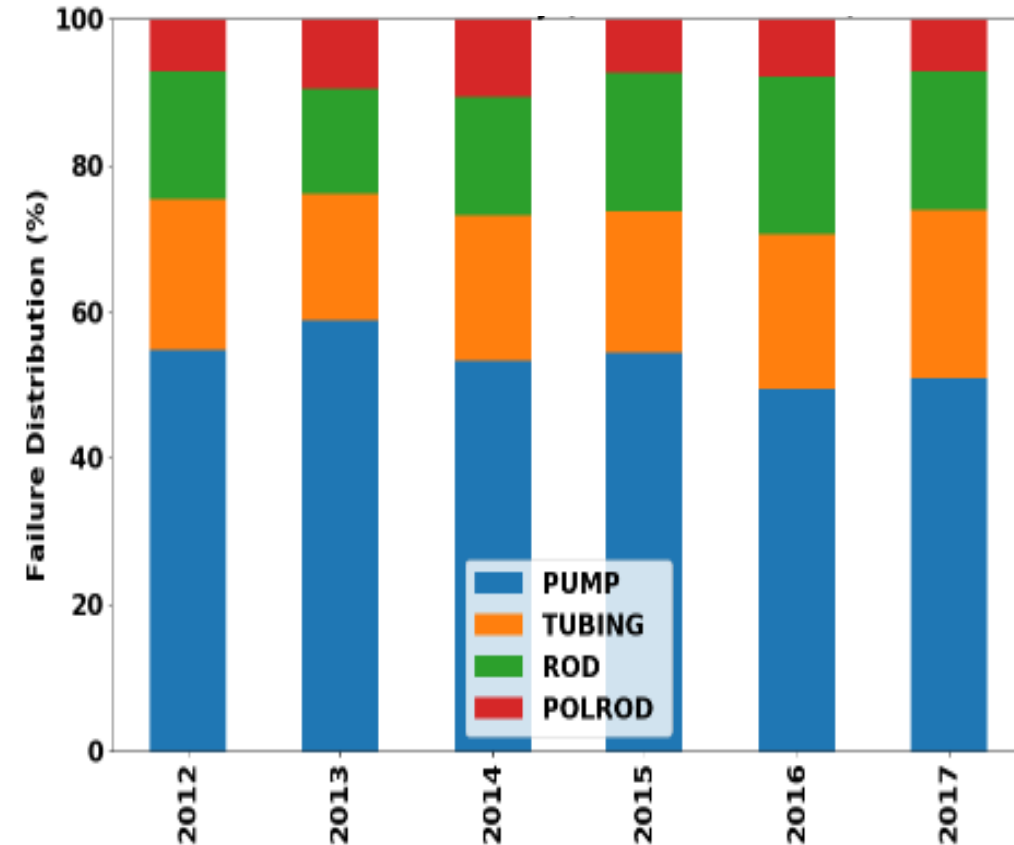
Dynocards decoded

Dynocard classifier

Failure Prediction

Understanding Data

- Well data collected by POC
 - **Sensors:** Load cells, transducers, etc.
 - **Data:** Surface card, card area, peak/min load, SPM, csg/tbg pressure, pump fillage, etc.
 - These signals are recorded in XSPOC databases
- Analog and dynamometer card data simultaneously available only from Dec 2017



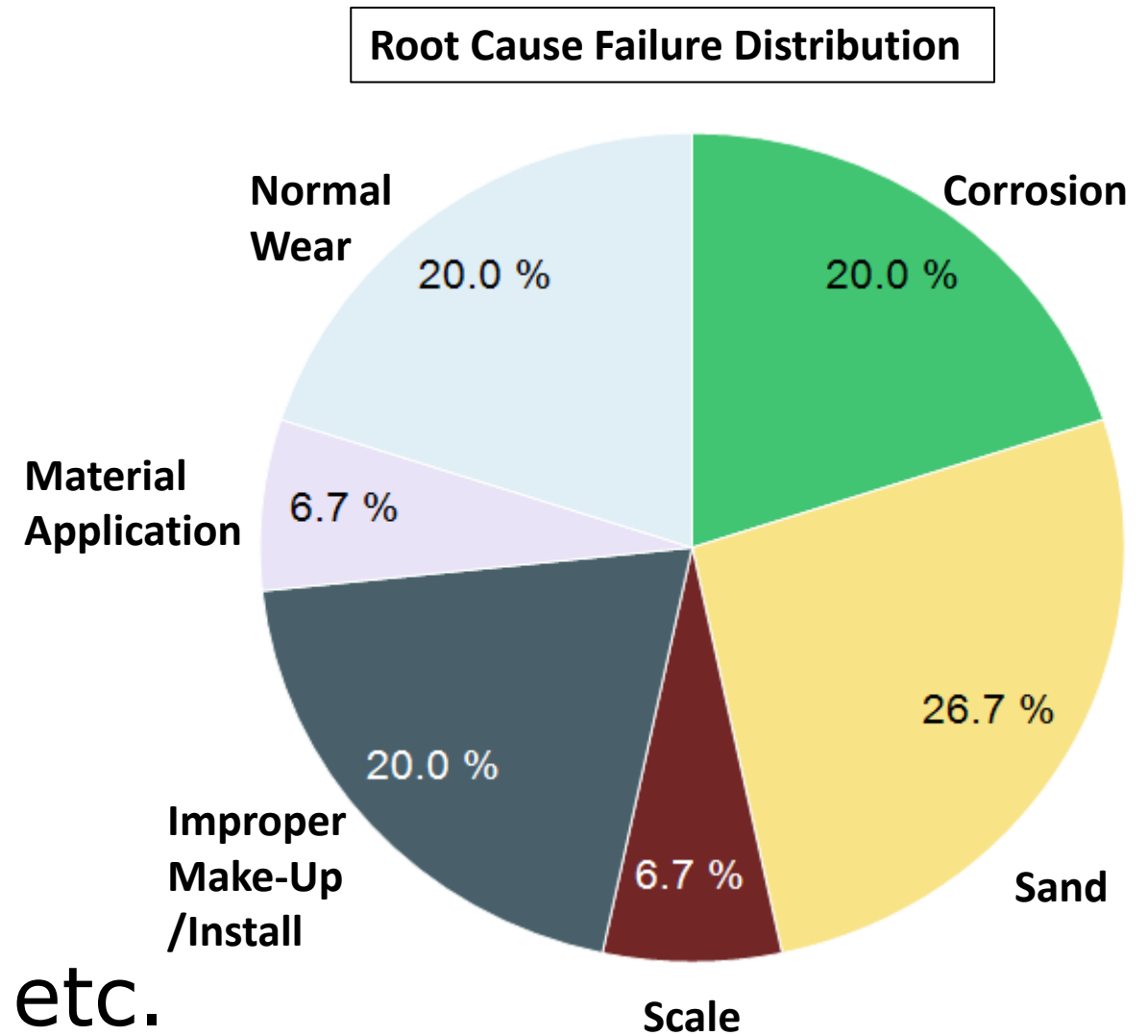
Failure Root Cause

- **Mechanical**

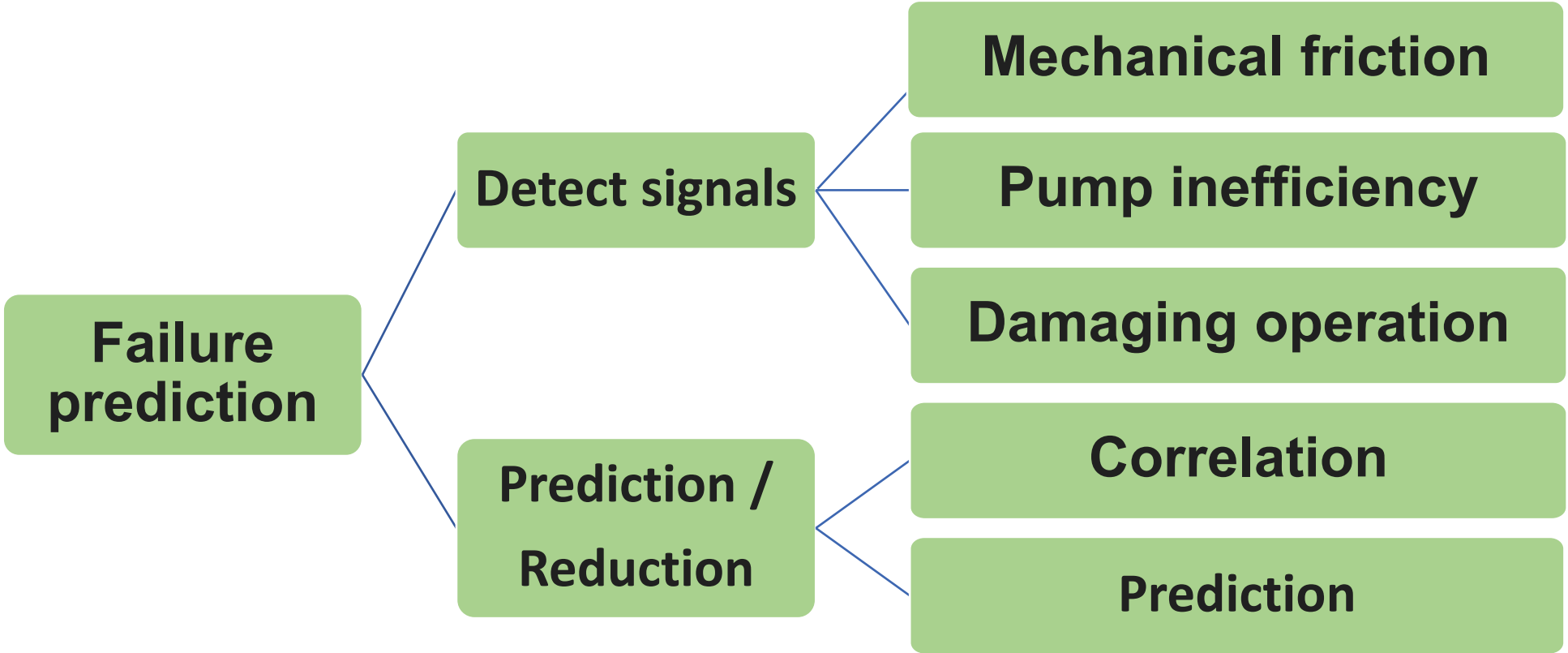
- Improper design/manufacturing
- Wear and tear during operations
 - Sand intrusions
 - Fluid pounding
 - Rod cutting
 - Asphaltting

- **Chemical**

- Corrosion by H₂S, CO₂, etc.



Applying Machine Learning to Rod Pumps



Dynamometer card

- Dynocard data is health indicator for a rod pump well
- More than **100,000 cards per day** stored by CRC wells
- Classification enables time-series visualization of card data
- Dynocard visualization over time facilitates suboptimal well diagnostics and failure prediction

Interpreting Dynamometer card



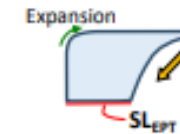
Ideal Card: fully anchored tbg, 100% liquid fillage, & pump in good condition.



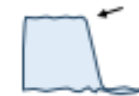
Slanted: Unanchored tbg indicated by the card being slanted at the k_{tbg} (Tubing Spring Constant).



Fluid Pound: sudden impact load. Inefficient and very damaging to pump, rods, tubing, and GBox. The impact load causes rod buckling & rod-on-tbg slap.



Gas Interference (or Gas Pound): a more gradual load transfer as gas compresses (pneumatic cushioning). Greatly reduces the pumping efficiency and indicates the well is not pumped off (\approx Fluid# @ a higher PIP).



Hole In Barrel: as the bottom of the plunger passes the hole (arrow) the hydrostatic pressure is equalized across the plunger causing the F_o to be lost.



Worn Pump: slow to pick up & quick to release the fluid load, due to: TV leaking or plunger/barrel wear.

Building a Dynocard Classifier

>100k cards

Feature engineering

- Card area
- Perimeter
- Area above card
- Area Below card
- Fillage
- Compression length
- Max load
- Max position
- Cumulative load
- Load center
- Position center

K-means Clustering



Clusters to Classes

Class #1: Full Card

Class #2: Pump off

Class #3: Fluid Pound (Medium)

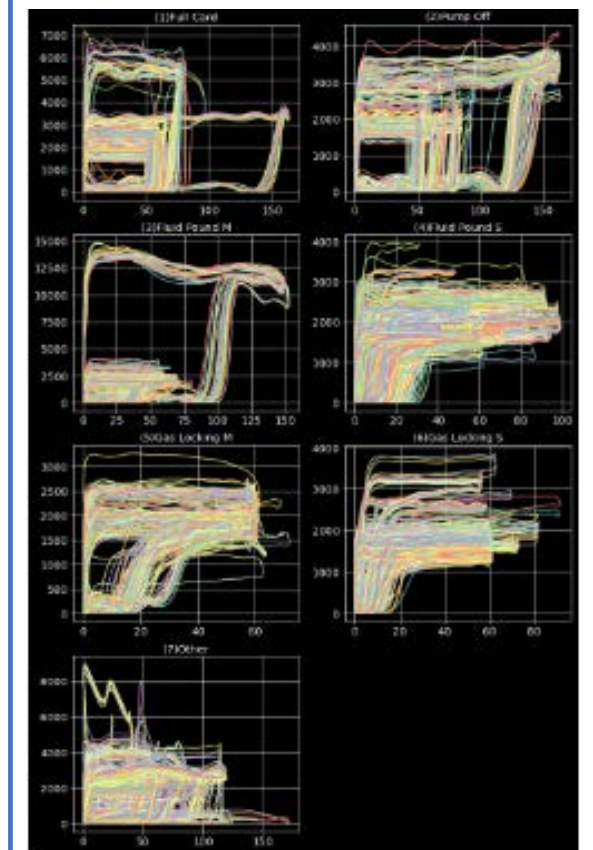
Class #4: Fluid Pound (Severe)

Class #5: Gas Interf.

Class #6: Gas Locking

Class #7: Other cards

Classifier Model



Well Failure Prediction: Industry Status

- Can we predict failures?
- USC-Chevron, PhD and MS students
(SPE 165374, **2013**)

	Tubing & Pump Failure
Precision (%)	65

- Ospreydata Inc, founded in 2013
(SPE 190090, **2018**)

	Tubing Failure	Pump Failure
Precision (%)	54	71

Rod Pump Failure Analytics

Diagnostic

When did it fail?

Failure Detection

Descriptive

Why did it fail?

Failure Correlation

Predictive

Is it going to fail?

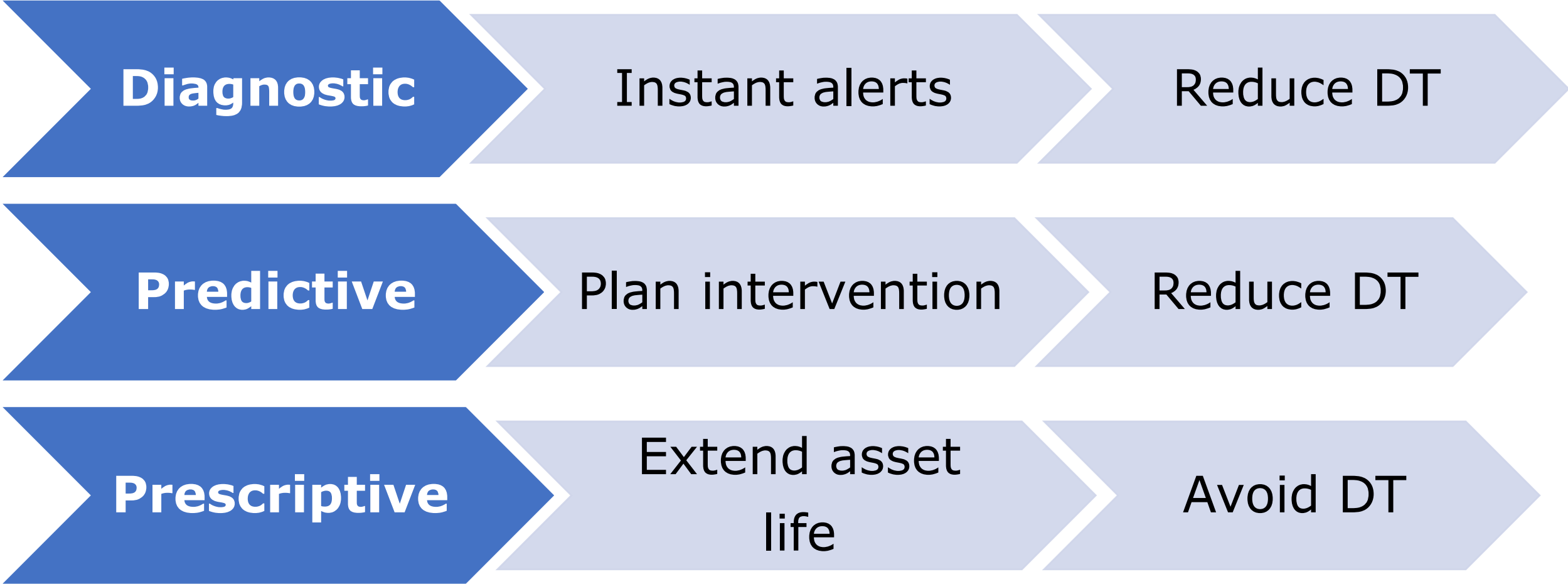
Failure Prediction

Prescriptive

How to prevent failure?

Failure Elimination

Value of Analytics



Failure Prediction

Pump



Tubing



Rod string

Workflow

Data pre-processing

Understand indicators of failures (signals and card classes)

Detect early signals leading to an event

Build database of historical pre-failure signals (manual vs automated)

Train predictive failure models

Maintain and improve models

Data Pre-processing Data

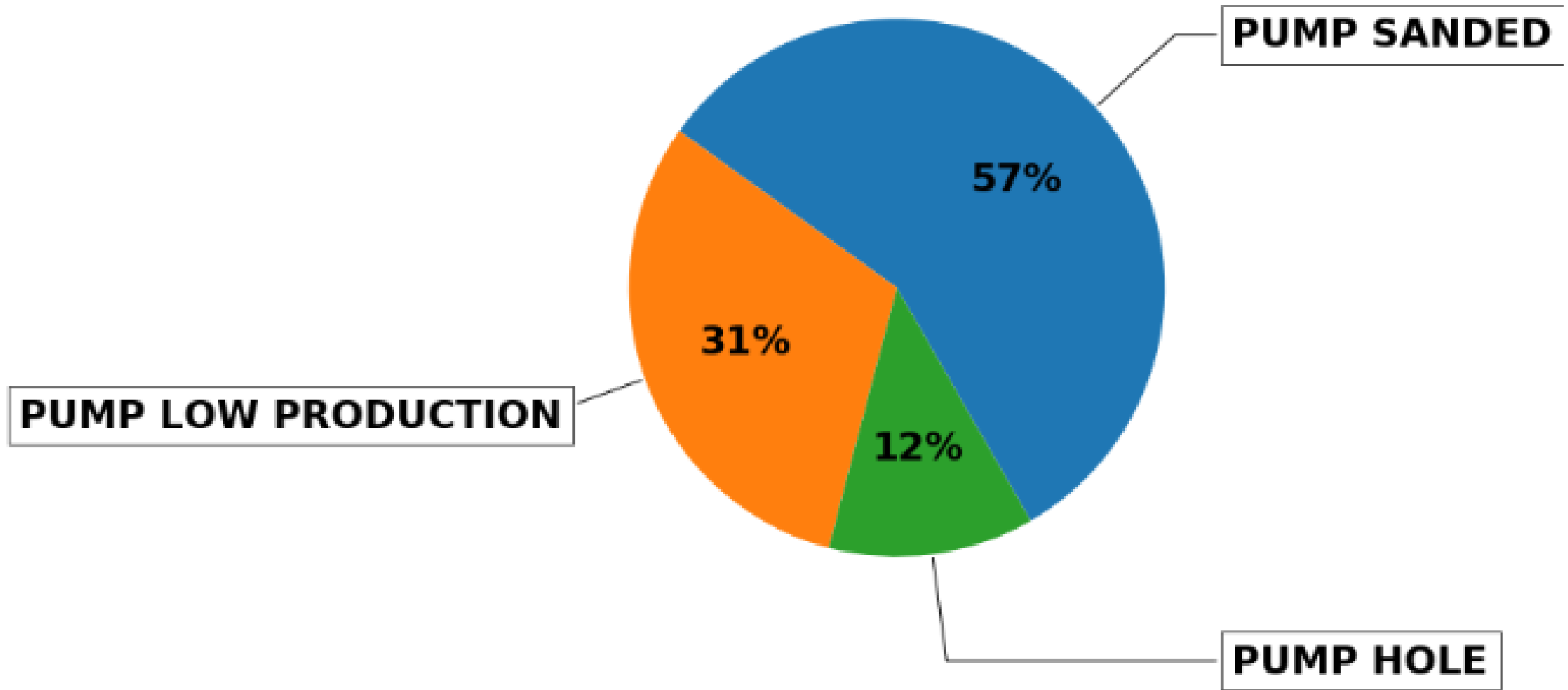
- Filter wells
 - Method of production: Rod pump
 - Primary failure mode, e.g. Pump
 - Secondary failure mode, e.g. Sanding, low production, scaling, etc.
- Consolidate analog data with dynocard features
 - Area, perimeter, centroids, area above, area below, etc.
- Unify time resolution. Interpolation needed?
 - Forward fill, backward fill, linear interpolation

Feature Engineering

- Original features: Peak Load, Min Load, Run time, area, fillage, dynocard features, etc.
- Normalize features
- FE level 1: Backwards sliding window feature generation¹
 - A: Long-term statistical summary (e.g. last month)
 - B: Short-term statistical summary (e.g. last week)
 - C: Current Statistical summary (e.g. last 3 days)
 - Feature group 1: B/A
 - Feature group 2: C/A
- FE level 2: Include last n days of feature groups at each example

¹SPE 165374

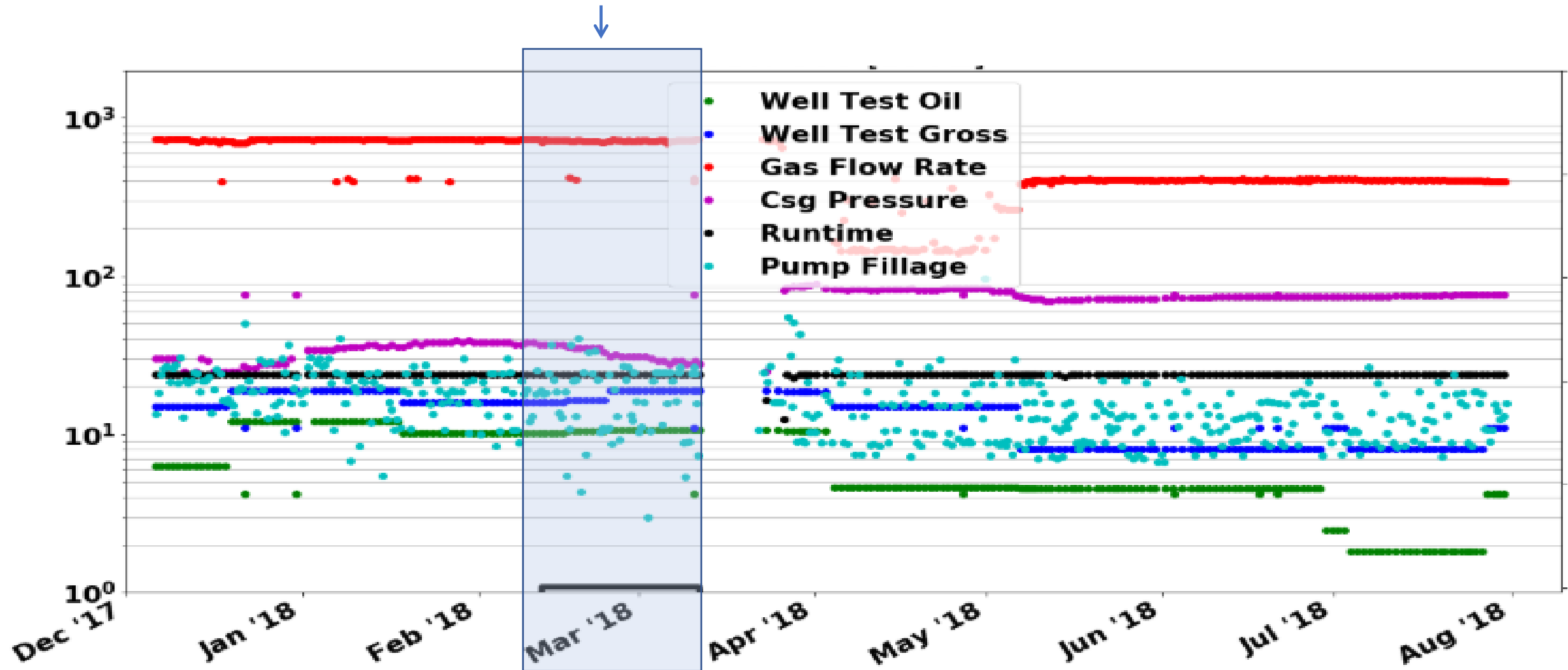
Pump Failure Modes/Root-Cause (Field X)



Pump Failure Due to Low Production

Pre-failure trend

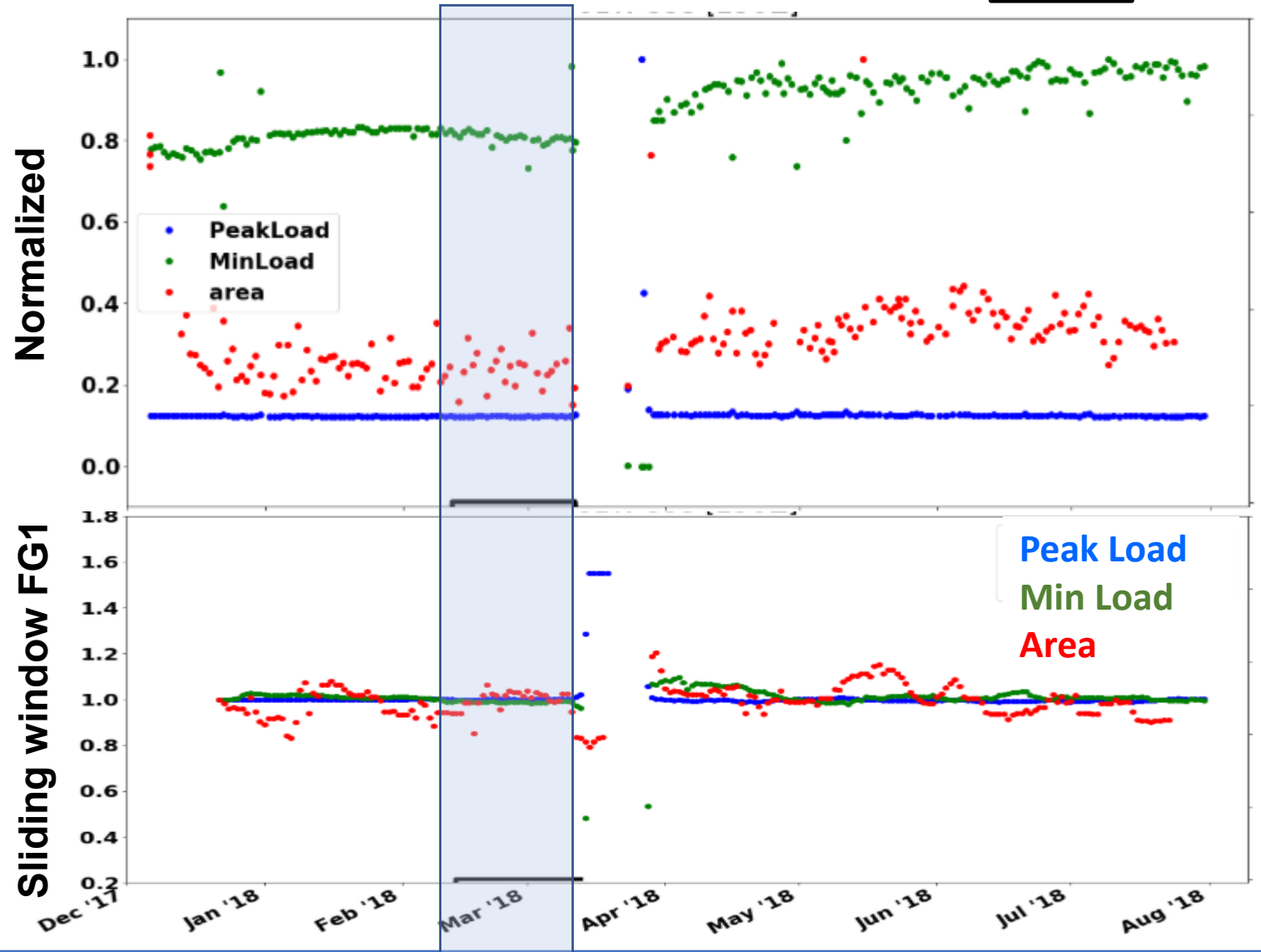
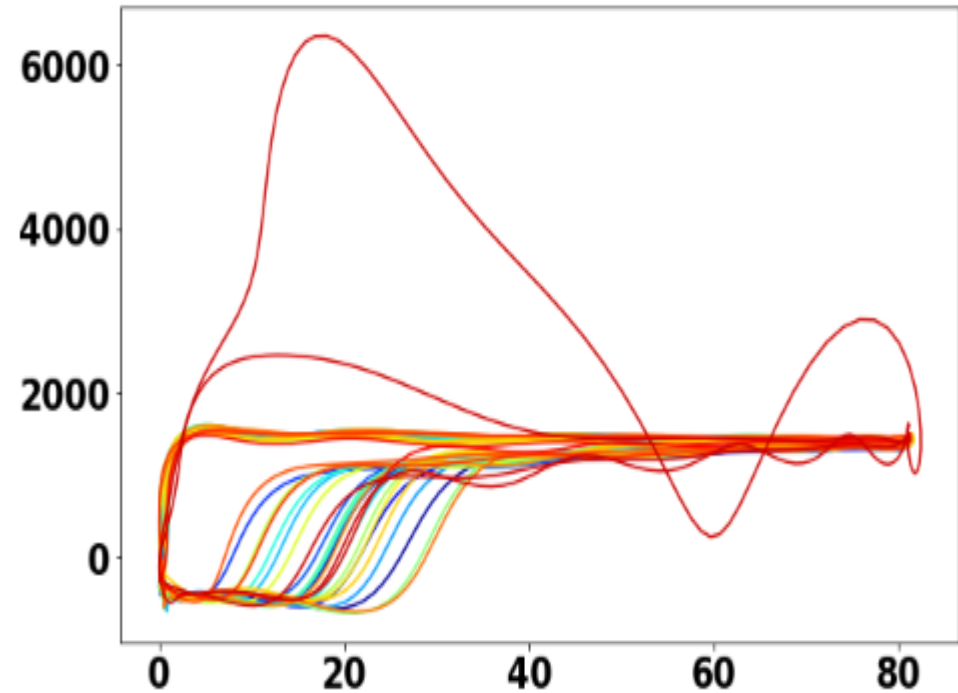
Well A



Pump Failure Due to Low Production

Well A

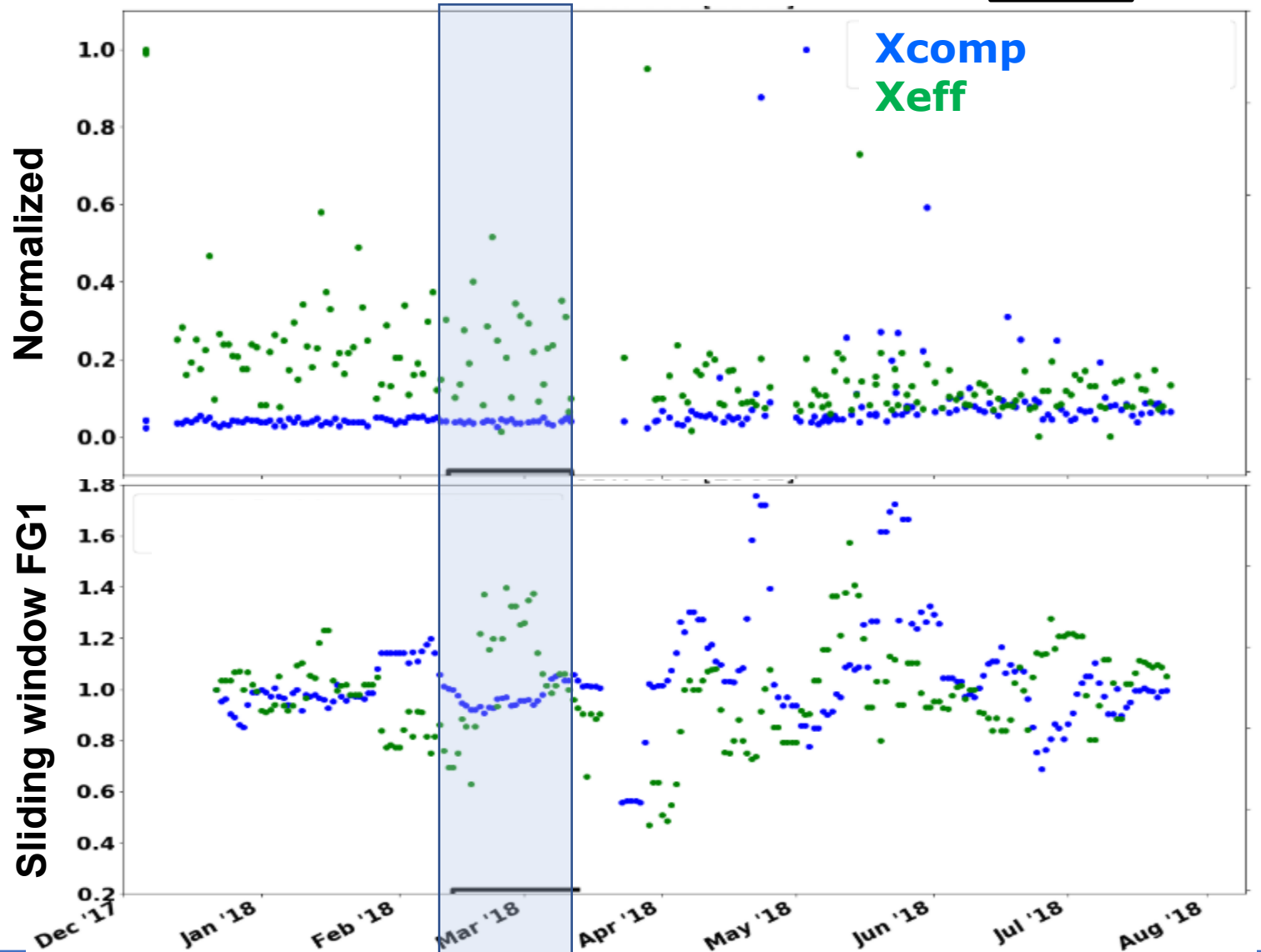
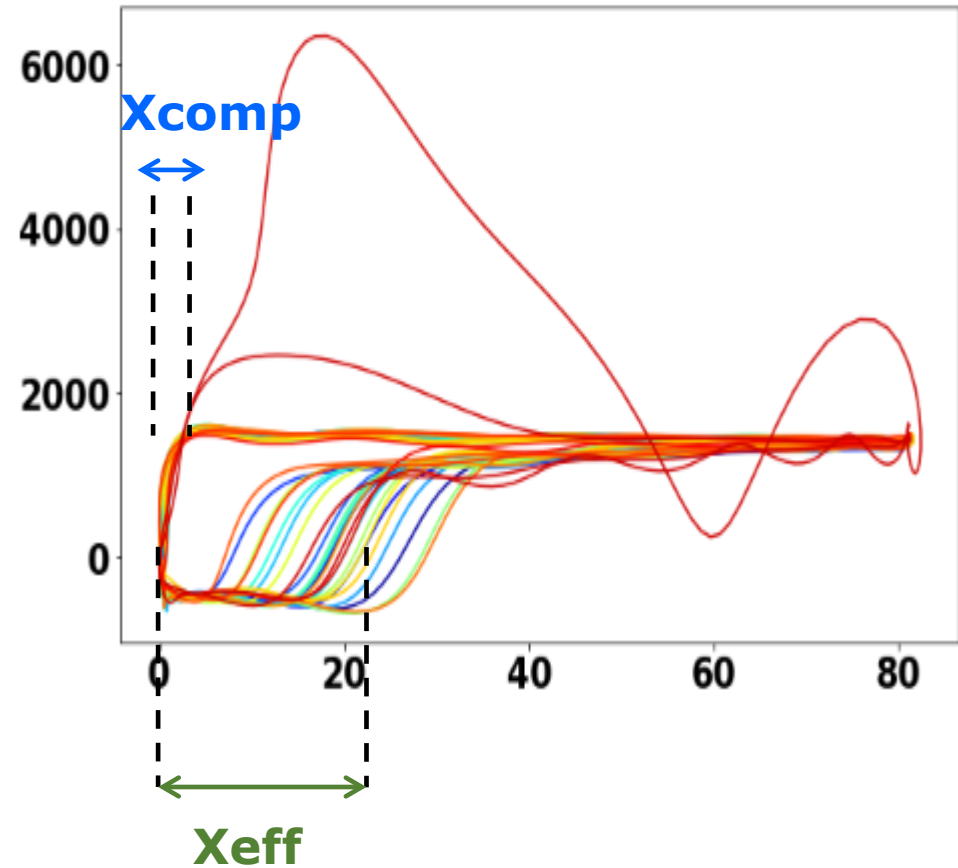
Blue (March 2018) -----> Red (Apr 2018)



Pump Failure Due to Low Production

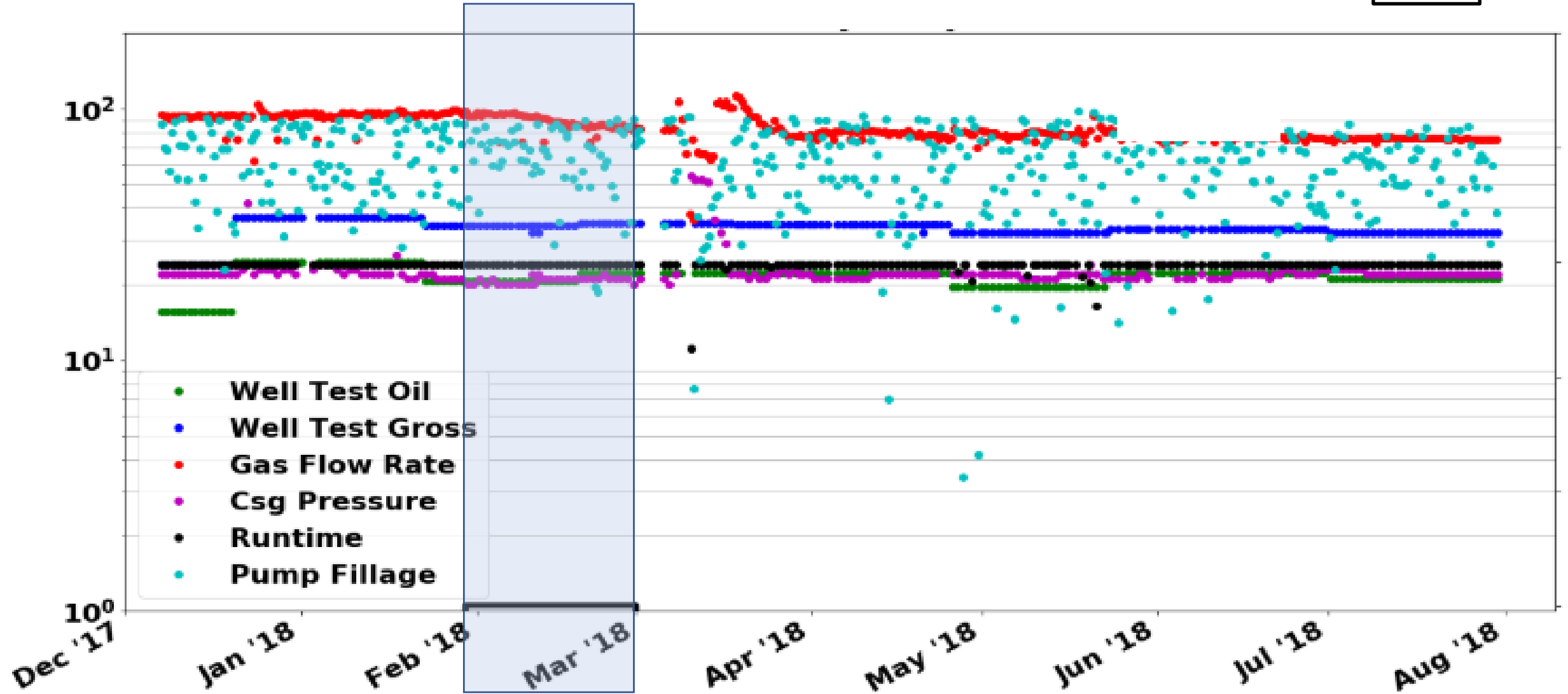
Well A

Blue (March 2018) -----> Red (Apr 2018)



Pump Failure Due to Sanding

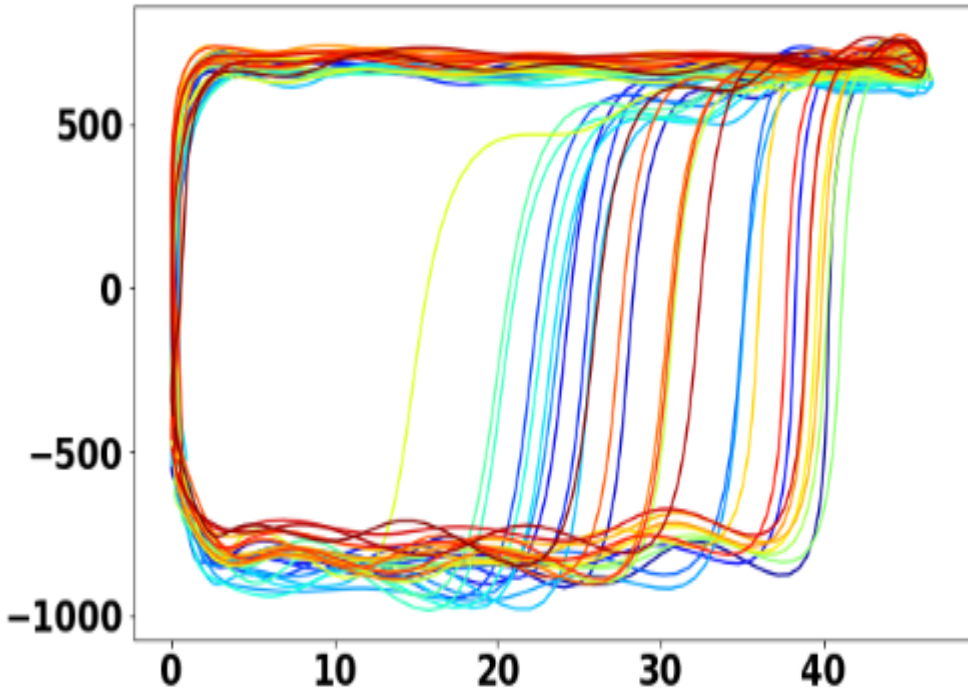
Well B



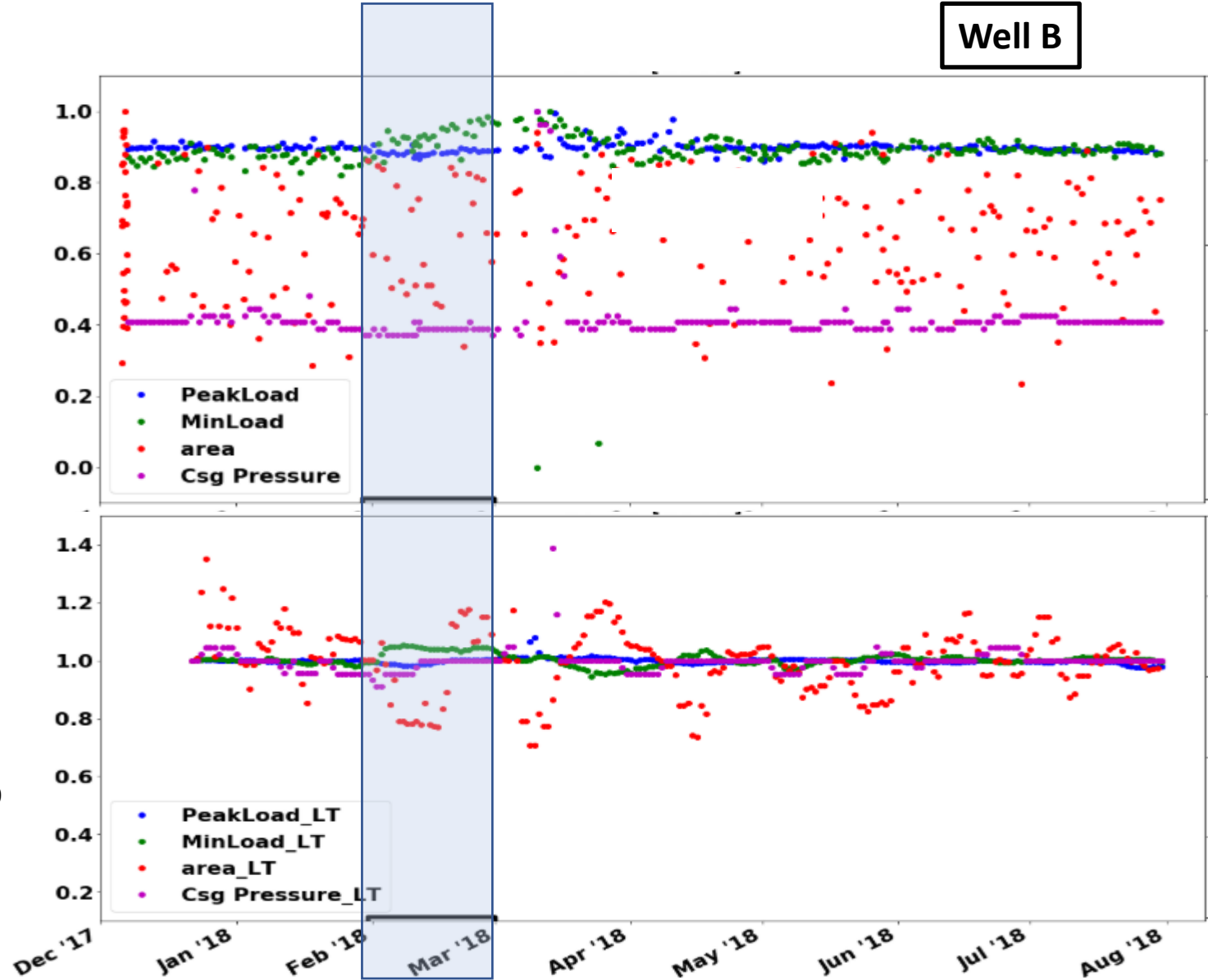
Pump Failure Due to Sanding

Well B

Blue (Feb.) -----> Red (Mar.)



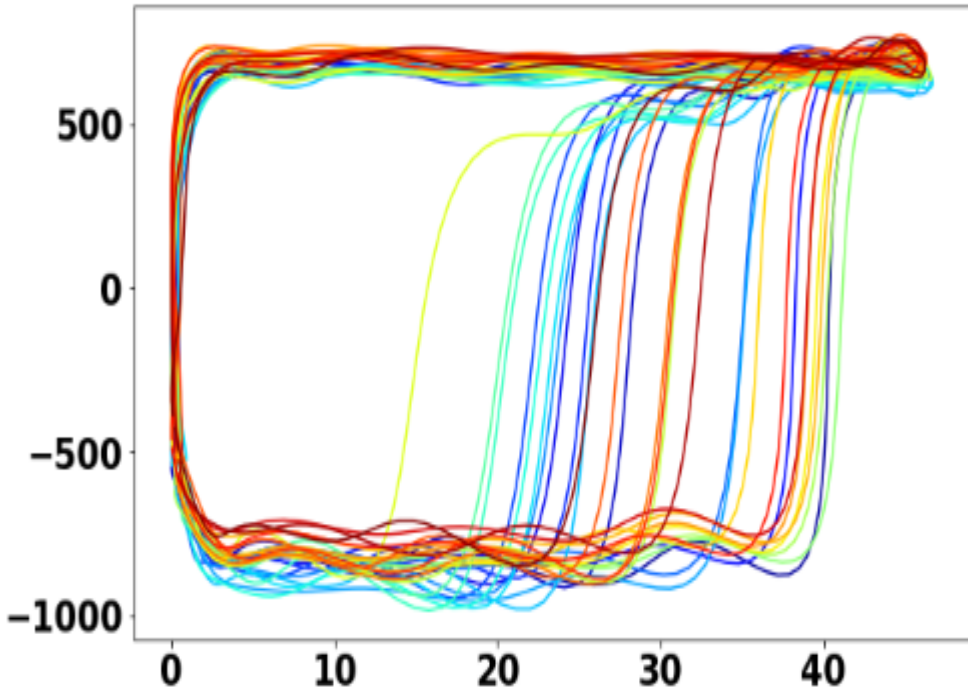
Normalized
Sliding window FG1



Pump Failure Due to Sanding

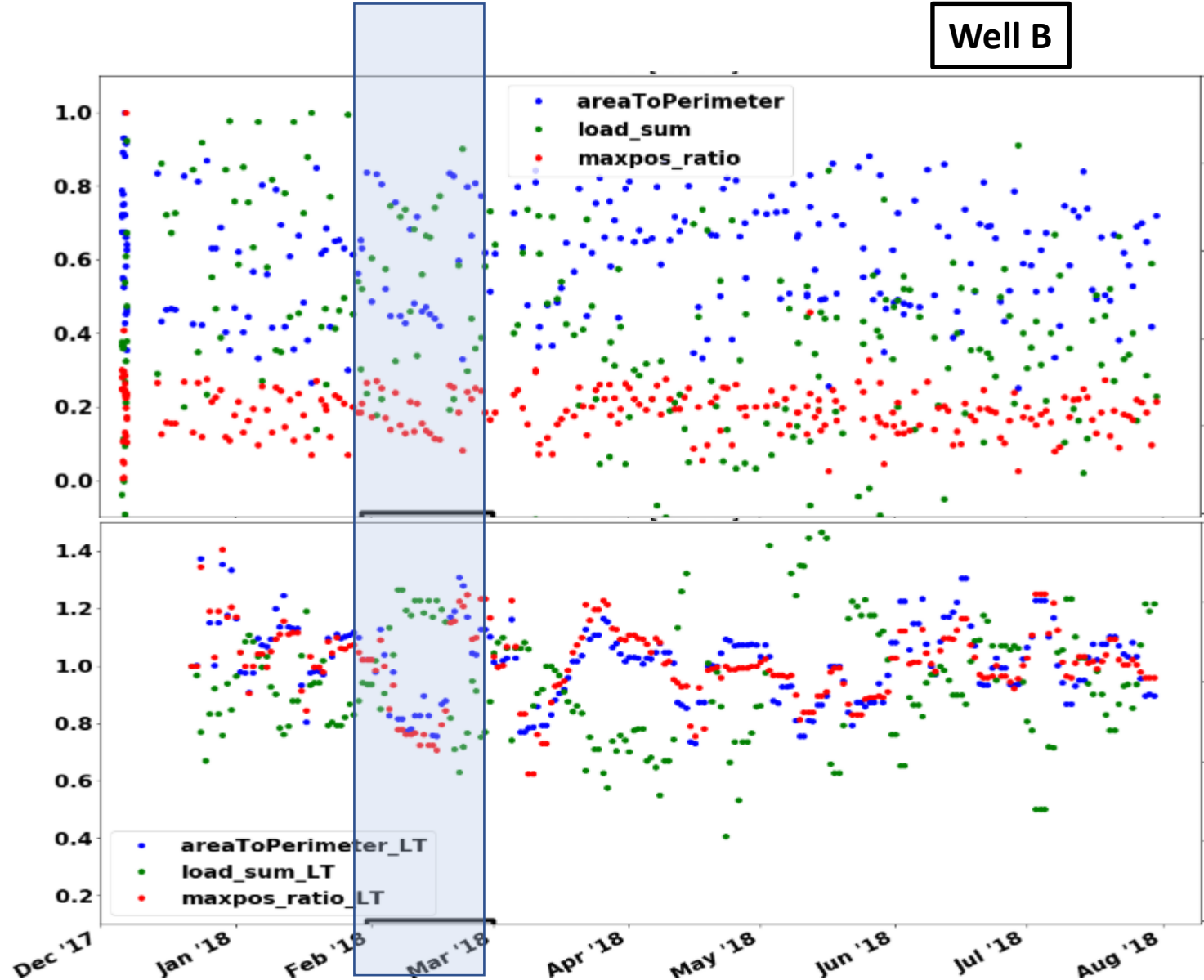
Well B

Blue (Feb.) -----> Red (Mar.)



Normalized

Sliding window FG1



Summary

- We can achieve full potential by preventive maintenance enabled by Rod pump prescriptive analytics.
- Prescriptive analytics requires failure predictive and descriptive capabilities.
- We are working with SMEs to more accurately determine failure intervals and label the root cause.
- Built dynocard classification model. We are working to use it for well optimization.
- We are expanding our failure data lake – card data are now permanently stored.

Thank You

Confusion Matrix and Precision-Recall Curve

Confusion Matrix

		Prediction	
		Positive	Negative
Actual	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Higher recall predicts **more true alarms**
But it may **produce a lot of false alarms**

$$Recall = \frac{TP}{TP + FN}$$

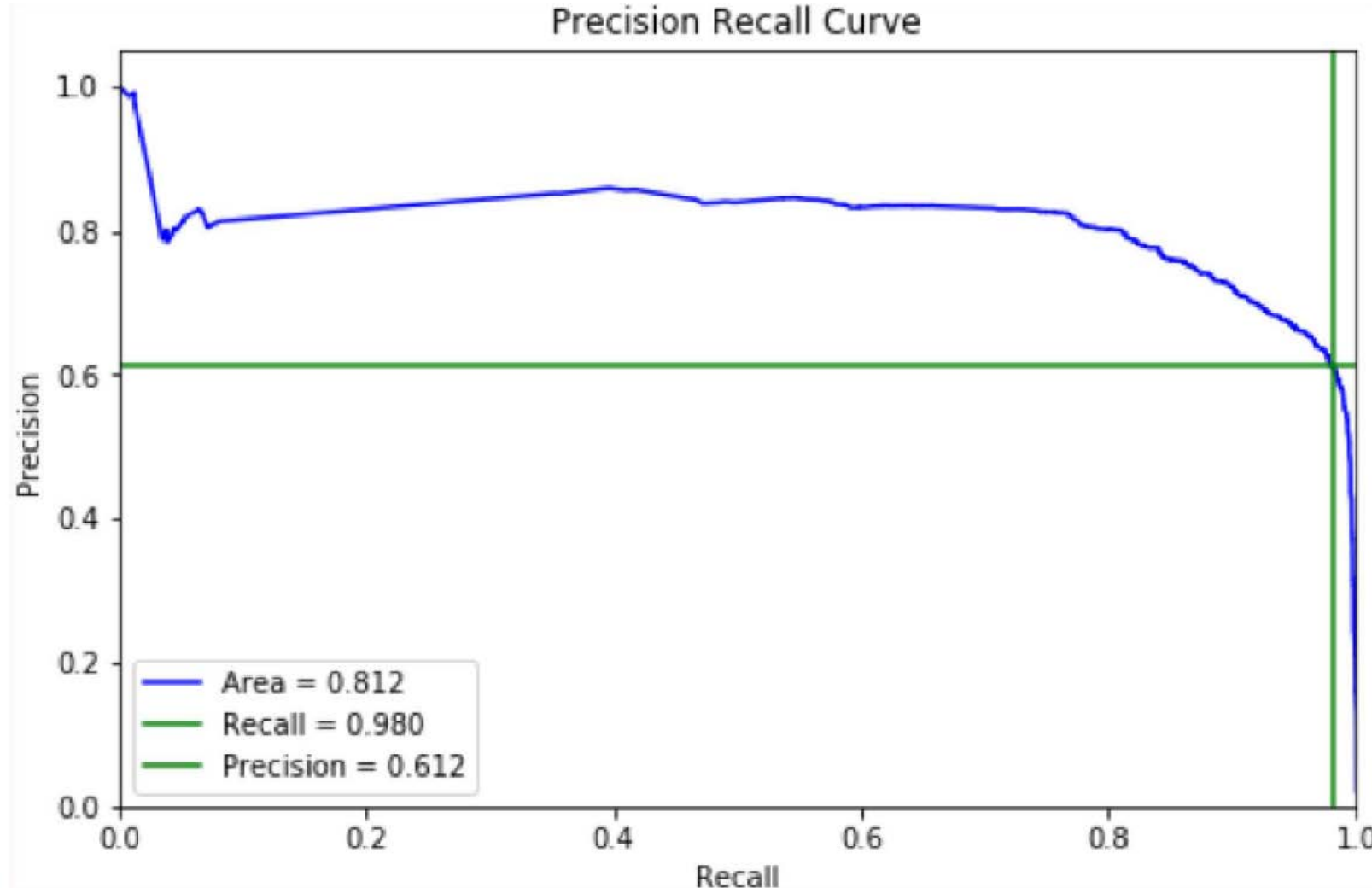
~~$Accuracy = \frac{TP + TN}{All}$~~

$$Precision = \frac{TP}{TP + FP}$$

Higher precision predicts **less false alarms.**
But we may **miss a lot of true alarms**

Precision-Recall Curve

- Model performance is adjustable by the business value



Ref: SPE 190090 [Tubing failure]