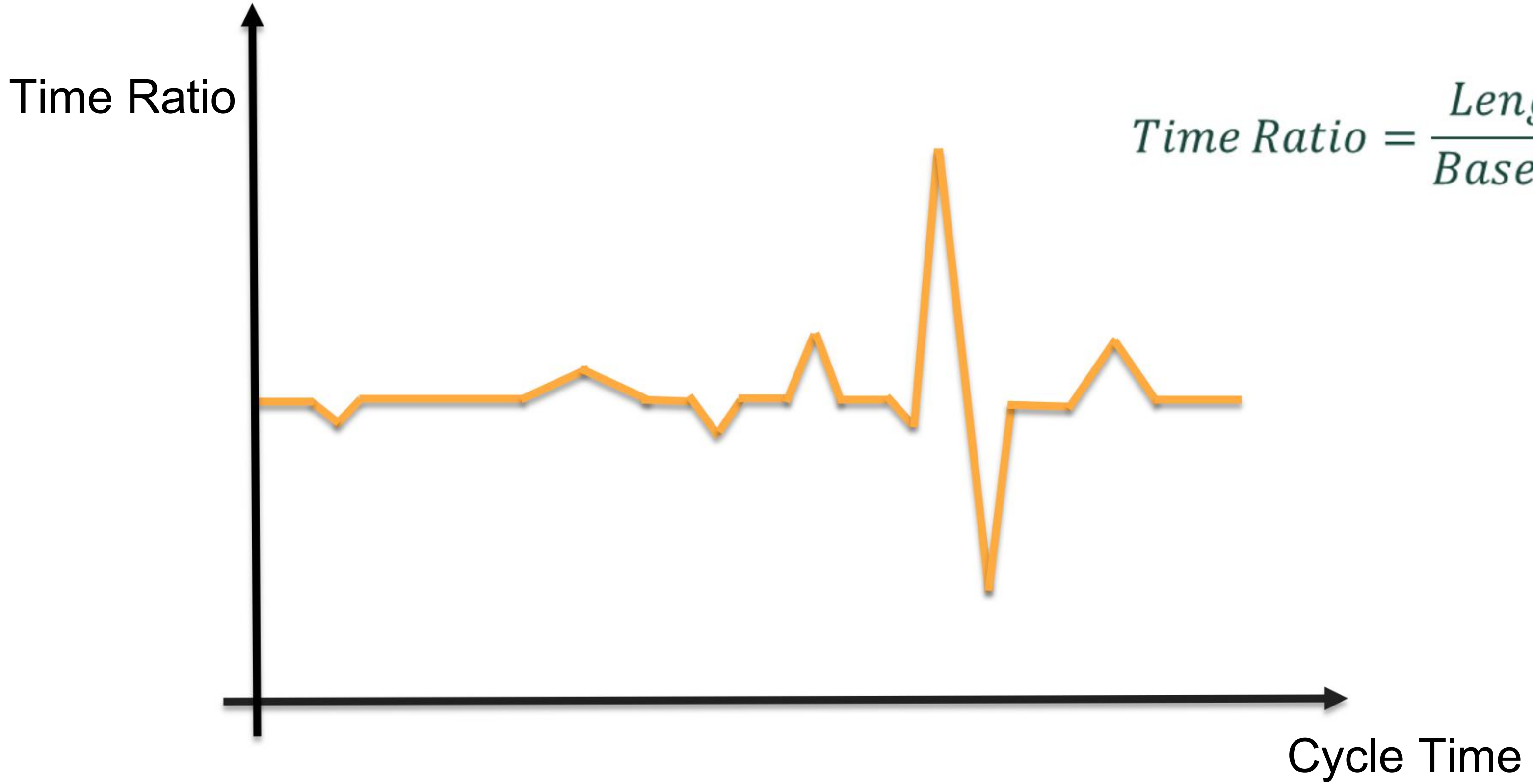
Sinto Project

Quan Xiao
9/20/2019

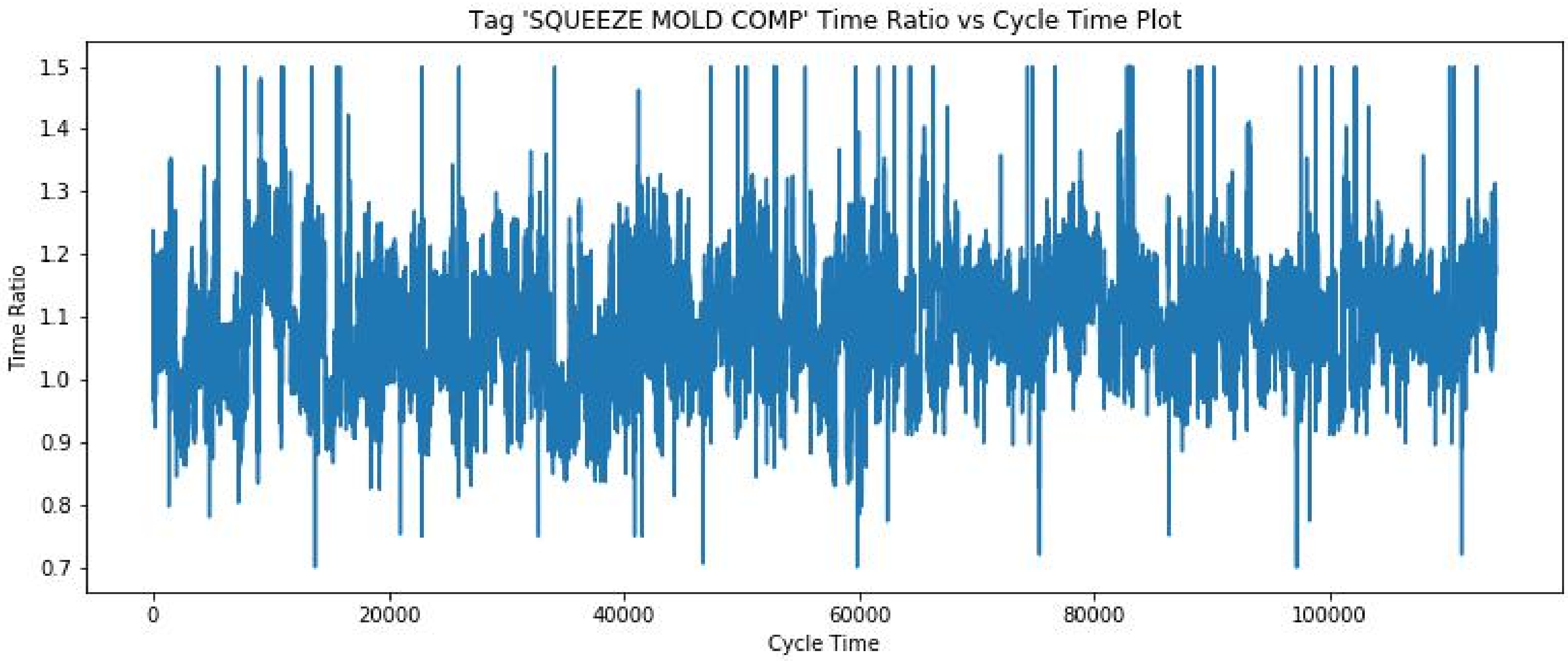
Tag	DateTime	Length	Base line	Model	Tag	DateTime	Length	Base line	Model
MOLD PUSH OUT RET SINGLE COMP.	8/28/2019 5:46	0	2.5	Single	MOLD PUSH OUT RET SINGLE COMP.	8/28/2019 5:46	0	2.5	HI-MED
CHAPLET SETTING COMPLET ED	8/28/2019 5:46	0	25	Single	CHAPLET SETTING COMPLETE D	8/28/2019 5:46	0	25	HI-MED
COPE PATTERN SPRAY COMP	8/28/2019 5:46	0.409	1	Single	COPE PATTERN SPRAY COMP	8/28/2019 5:46	0.41	1	HI-MED
DRAG PATTERN SPRAY COMP	8/28/2019 5:46	0.515	1	Single	DRAG PATTERN SPRAY COMP	8/28/2019 5:46	0.513	1	HI-MED
INFLATE GATE SEAL COMP	8/28/2019 5:46	0.534	0.5	Single	INFLATE GATE SEAL COMP	8/28/2019 5:46	0.521	0.5	HI-MED
STRIP FLASK DOWN SET COM P	8/28/2019 5:46	0.02	1	Single	STRIP FLASK DOWN SET COM P	8/28/2019 5:46	0.02	0.1	HI-MED
SQUEEZE ADV SET COMP	8/28/2019 5:46	1.201	1.7	Single	SQUEEZE ADV SET COMP	8/28/2019 5:46	1.209	1.5	HI-MED
MOLD PUSH OUT ADV COMP	8/28/2019 5:46	5.928	6.6	Single	MOLD PUSH OUT ADV COMP	8/28/2019 5:46	5.934	6.6	HI-MED
...	...	...	...	...	...	...	...	...	...
COPE STOPPER OUT COMP	8/28/2019 6:11	0.38	0.5	Single	DRAG FLASK SLD RET COMP	8/28/2019 5:46	1.583	2	HI-MED
DRAG FLASK SLD RET COMP	8/28/2019 6:11	1.591	2	Single	SAND FILL COMP.	8/28/2019 5:46	14.535	15	HI-MED
COPE FLASK LIFTER DWN FO R STRIP COMP	8/28/2019 6:11	1.512	1.5	Single	COPE FLASK LIFTER DWN FO R STRIP COMP	8/28/2019 5:46	1.453	1.5	HI-MED
MOLD RECEIVER UP COMP	8/28/2019 6:11	2.853	2.6	Single	MOLD RECEIVER UP COMP	8/28/2019 5:46	2.854	2.6	HI-MED
STRIP FLASK DOWN STRIP	8/28/2019 6:11	2.012	1.7	Single	STRIP FLASK DOWN STRIP	8/28/2019 5:46	2.011	2	HI-MED
MOLD RECEIVER DOWN COM P	8/28/2019 6:11	1.503	1.2	Single	MOLD RECEIVER DOWN COMP	8/28/2019 5:46	1.451	1.2	HI-MED
CYCLE TIME	8/28/2019 6:11	1515.737	34.5	Single	CYCLE TIME	8/28/2019 5:46	32.836	24.2 5	HI-MED

We can gain several time series according to Tags



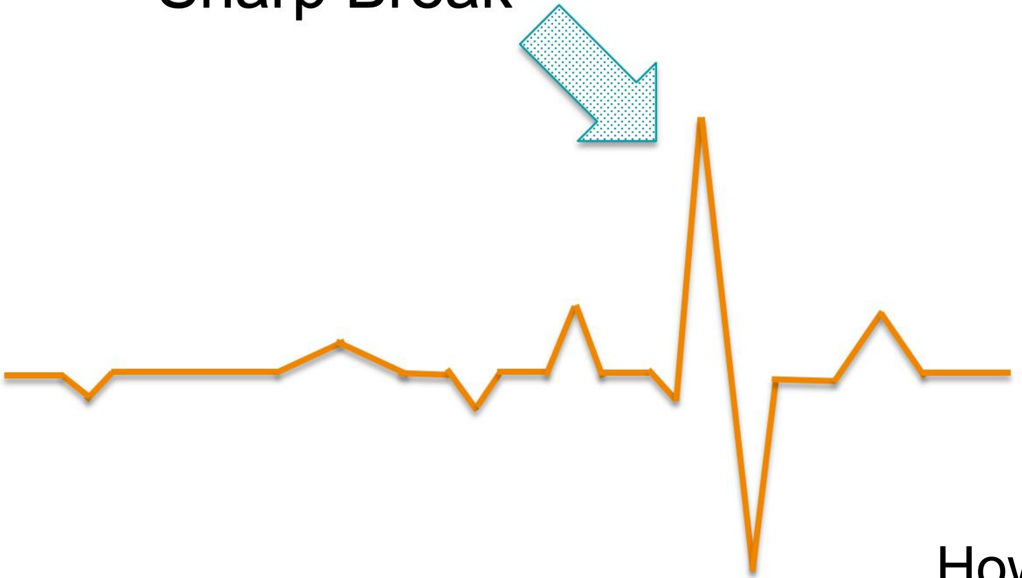
$$Time\ Ratio = \frac{Length}{Baseline}$$

The time series plot of Tag 'SQUEEZE MOLD COMP'



Note: The value of this plot is restricted from 0.7 to 1.5.

Sharp Break

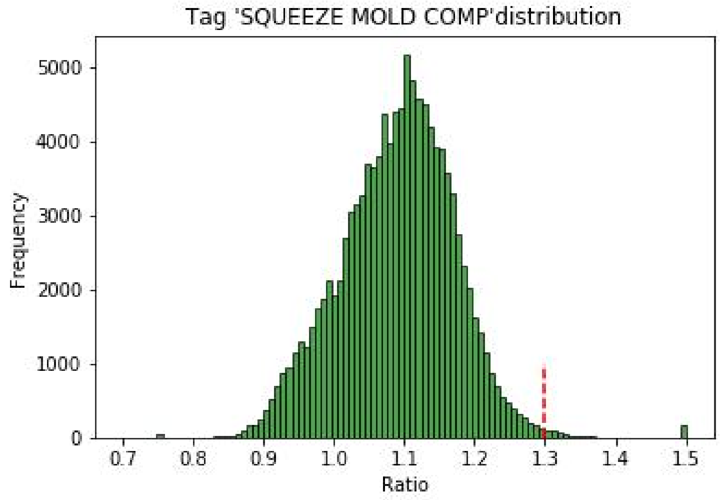


We want to prevent this breakdown before it happens.

How to do this?

The first thing we need to is to **define a threshold for outliers.**

- In order to do this, we draw the distribution of each Tag, eg. Tag 'SQUEEZE MOLD COMP'



- Since $\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.9973$

then we use $\mu + 3\sigma$ to be our threshold.
i.e. 1.3 in this Tag.

Related Problems

Problem Set-up

- With thresholds for anomalies => the labels => Classification.

Challenges:

- Unbalanced Data: The positively labeled data are less than 1%.
- In this situation, if we still use classification techniques, it's difficult to learn the positive labeled data well.

Anomaly Detection

Unsupervised learning:

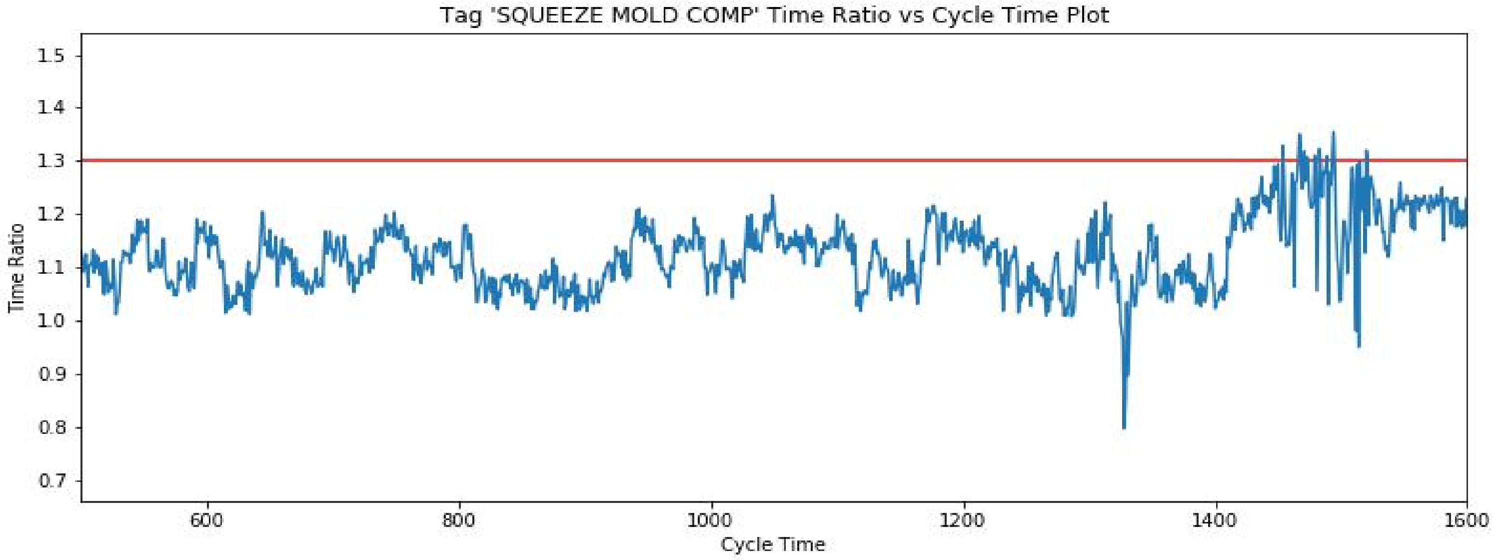
- Dataset is extremely unbalanced => only use normal data points to learn the normal patterns and reconstruct them.
- Feed another test data and calculate reconstruction error.
- Large error = Anomaly

Caveat:

- This detection technique can only tell us whether the points are outliers rather than predict them.

Proposed Method

- The pattern before the outlier is clearly different from the patterns before the normal points => Signal
- **How to verify?**



Problem Setup

Goal: Predict anomaly before it occurs.

Verify the areas before anomalies are also abnormal

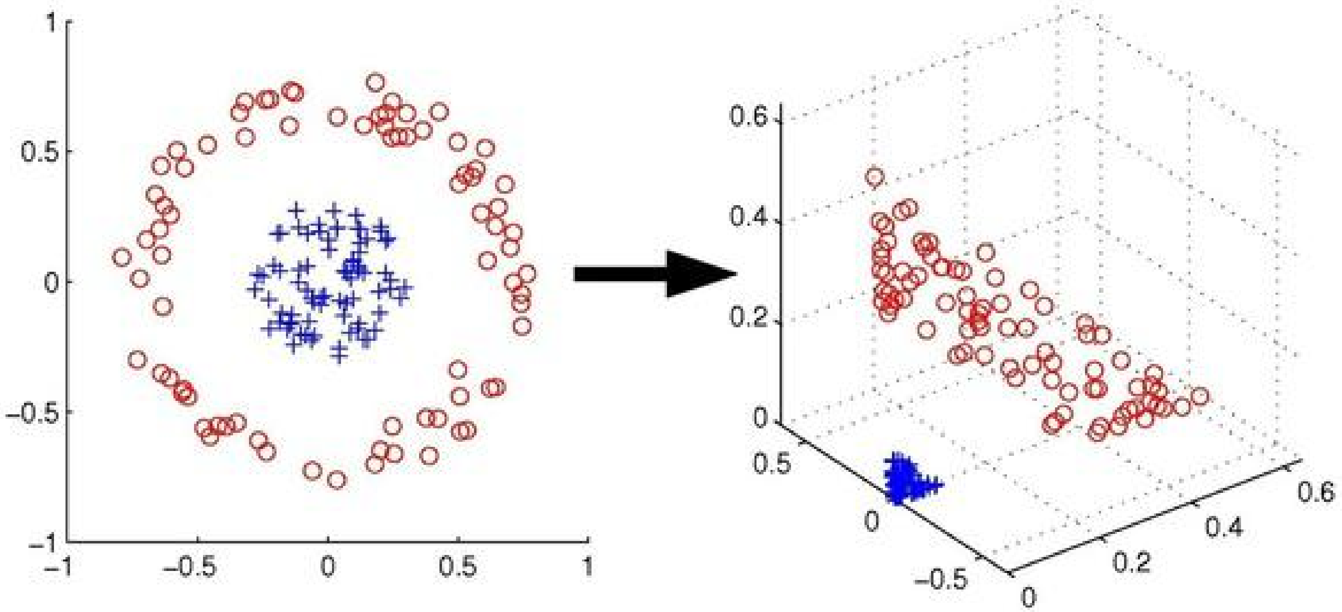
Predict the anomalies → Detect the area before anomalies

- Baseline model: Verify the feasibility of this hypothesis.

Baseline Model

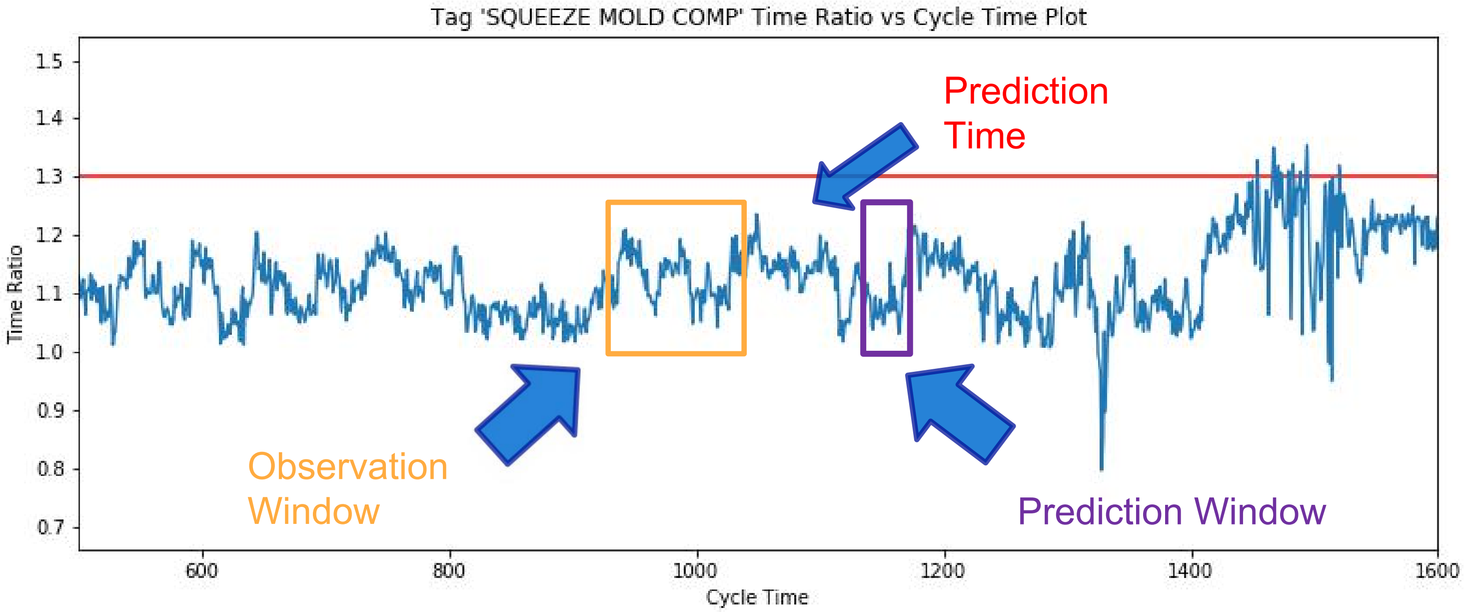
Baseline Model

kernel SVM (Support Vector Machine)



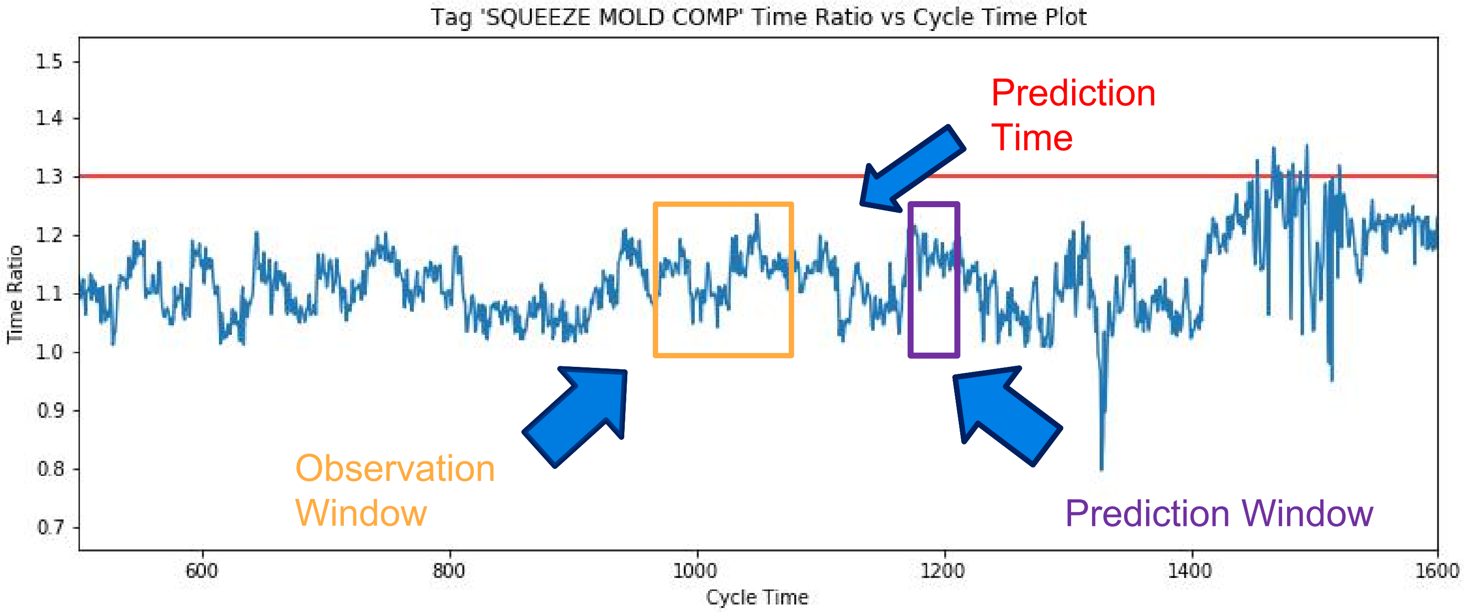
Intuition

- We use the features of observational Windows to predict if there are outliers in prediction windows.



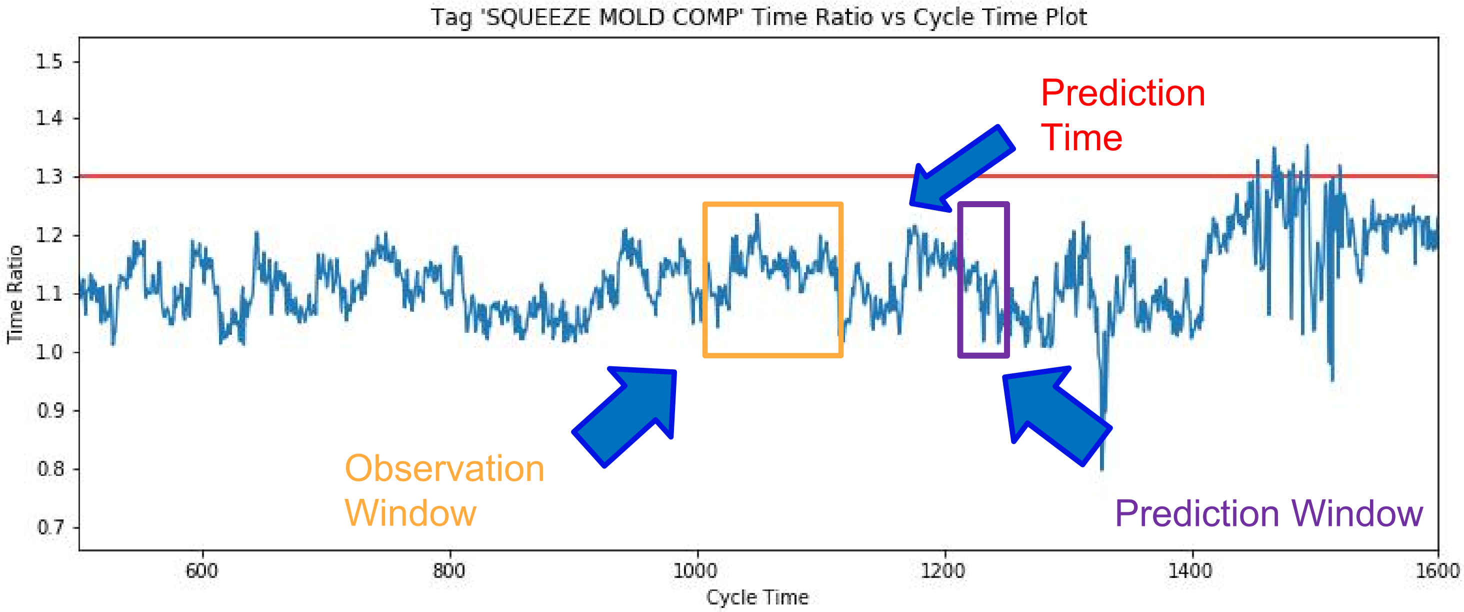
Intuition

- We use the features of observational Windows to predict if there are outliers in prediction windows.



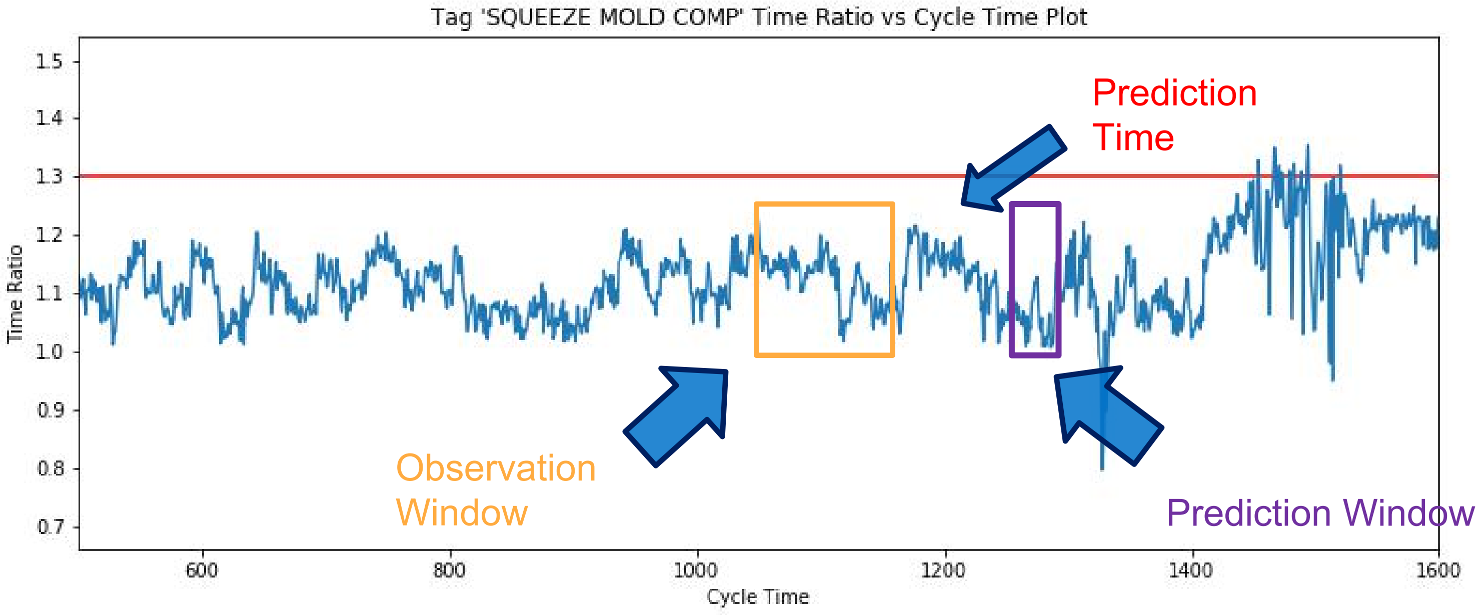
Intuition

- We use the features of observational Windows to predict if there are outliers in prediction windows.



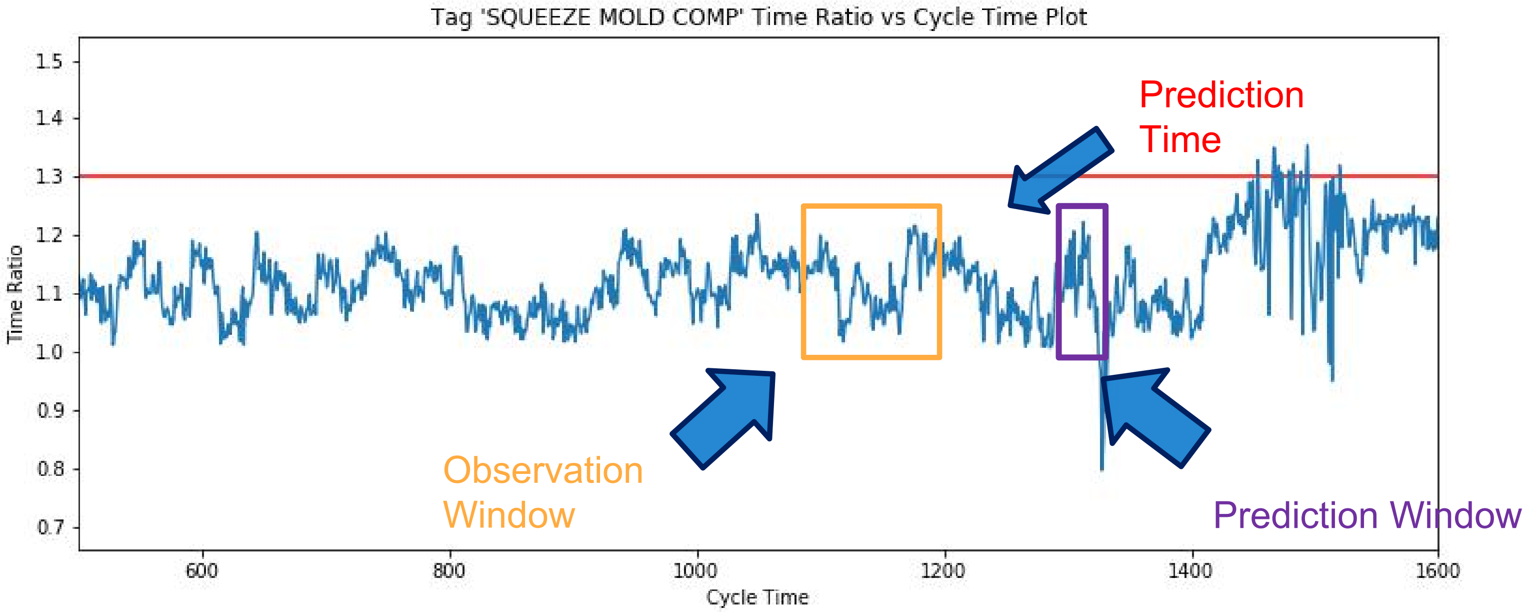
Intuition

- We use the features of observational Windows to predict if there are outliers in prediction windows.



Intuition

- We use the features of observational Windows to predict if there are outliers in prediction windows.

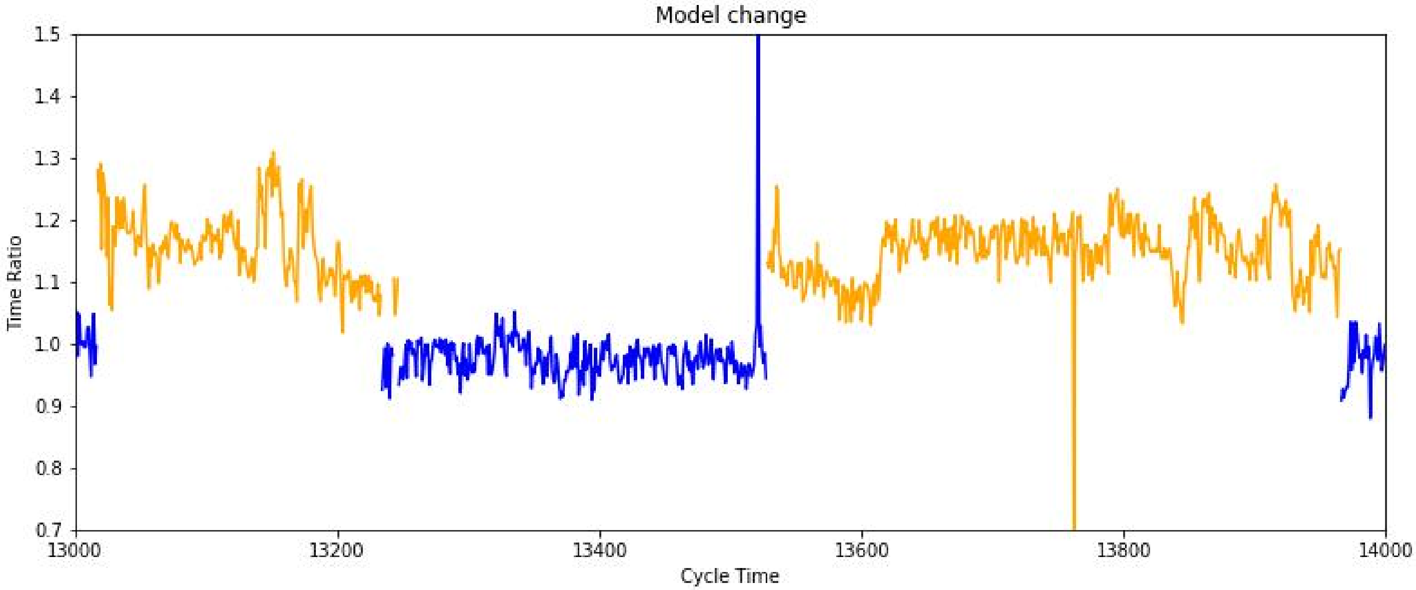


Features and Description

Feature	Description
Mean	The DC component (average value) of the signal over the window
Median	The median signal value over the window
Standard Deviation	Measure of the spreadness of the signal over the window
Variance	The square of standard deviation
Root Mean Square	The quadratic mean value of the signal over the window
Averaged derivatives	The mean value of the first order derivatives of the signal over the window
Skewness	The degree of asymmetry of the sensor signal distribution
Kurtosis	The degree of peakedness of the sensor signal distribution
Interquartile Range	Measure of the statistical dispersion, being equal to the difference between the 75th and the 25th percentiles of the signal over the window
Mean Crossing Rate	The total number of times the signal changes from below average to above average or vice versa normalized by the window length
Spectral Entropy	Measure of the distribution of frequency components
Model (one-hot encoded)	The model which the signal over the window belong to

Model Change

- We add Model types as features since the time ratio changes largely when change models



Feature Selection Method

- Why perform feature selection?
- Data interpretation discovery (insights into which factors which are most representative of our problem)

- Method: Sequential Forward Selection
- Greedy search algorithm

- How to assess the performance?

Selection criterion

	true condition	
Prediction condition	True positive (TP)	False positive (FP)
	False negative (FN)	True negative (TN)

$$\text{Precision} = \frac{tp}{tp + fp} \qquad \text{Recall} = \frac{tp}{tp + fn}$$

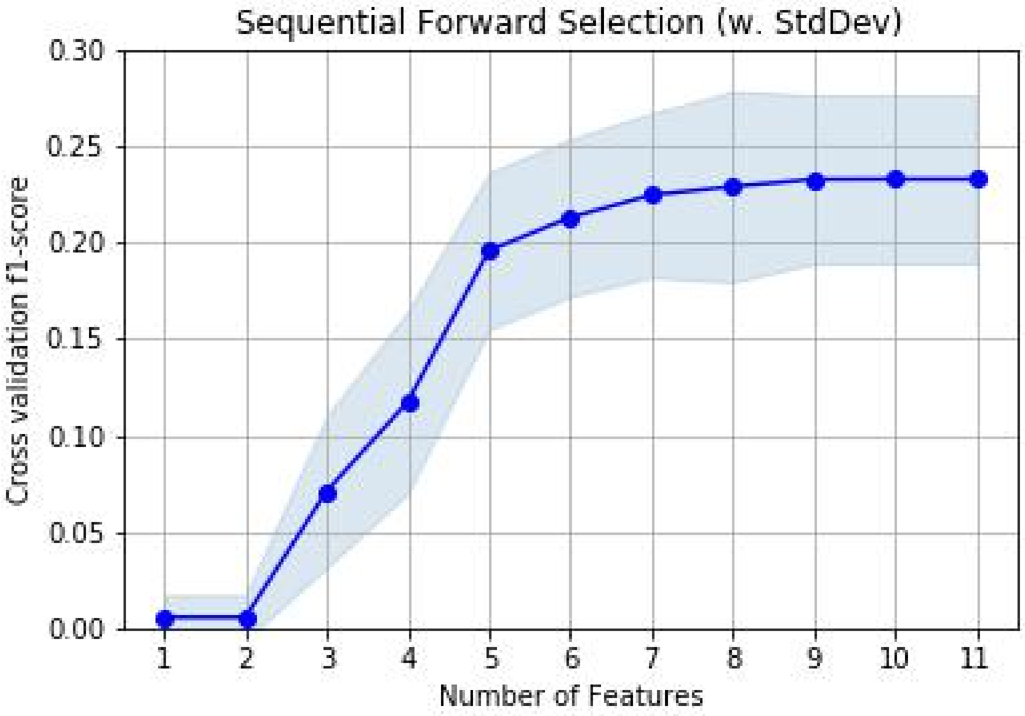
- We use F1 score which is defined as follows to be our selection criterion since it earns a trade-off between precision and recall.

$$\mathrm{F}_1 \cdot = \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}}$$

Selection Results

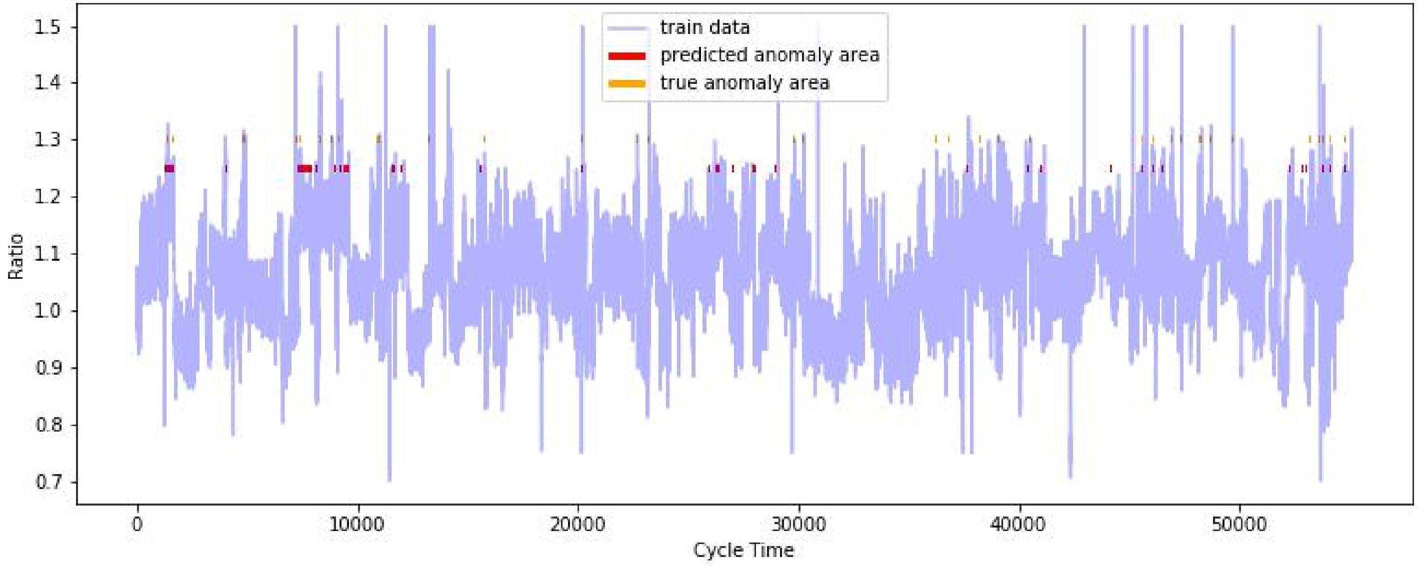
Selected Features:

- Mean
- Standard Deviation
- Root Mean Square
- Kurtosis
- Mean Crossing Rate
- Model 0,1,3,4,7,10



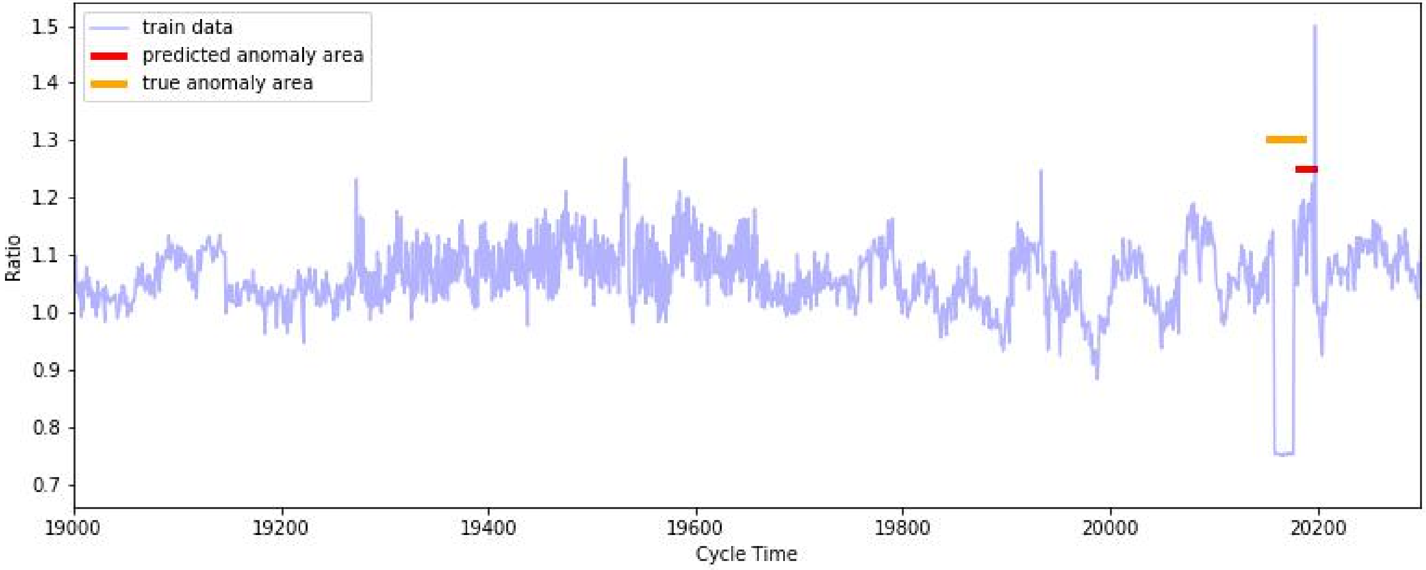
Fitting Result

- Train set



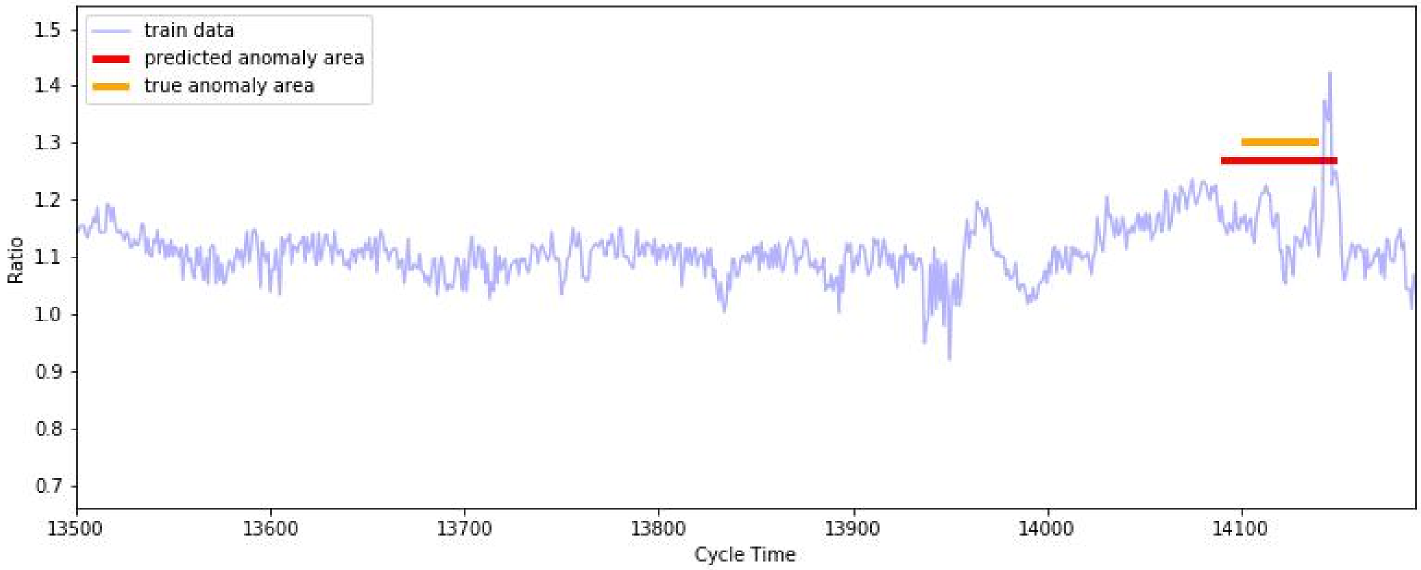
Fitting Result (Zoom in)

- Train set



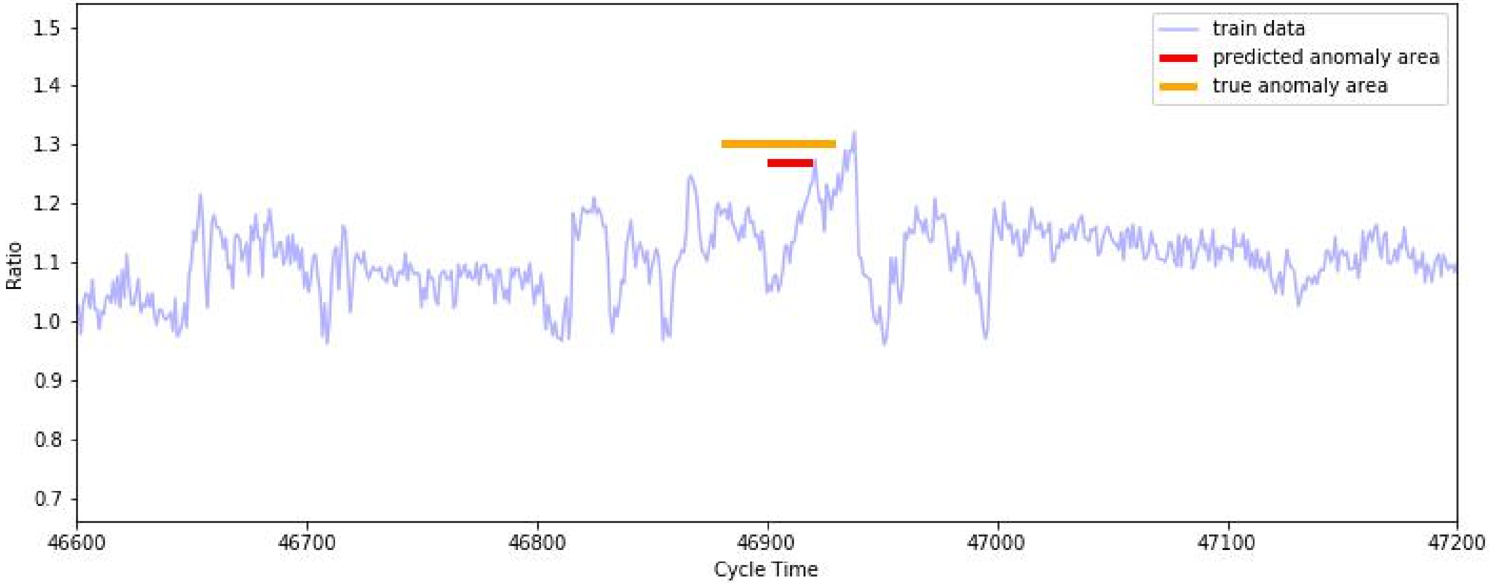
Fitting Result (Zoom in)

- Train set



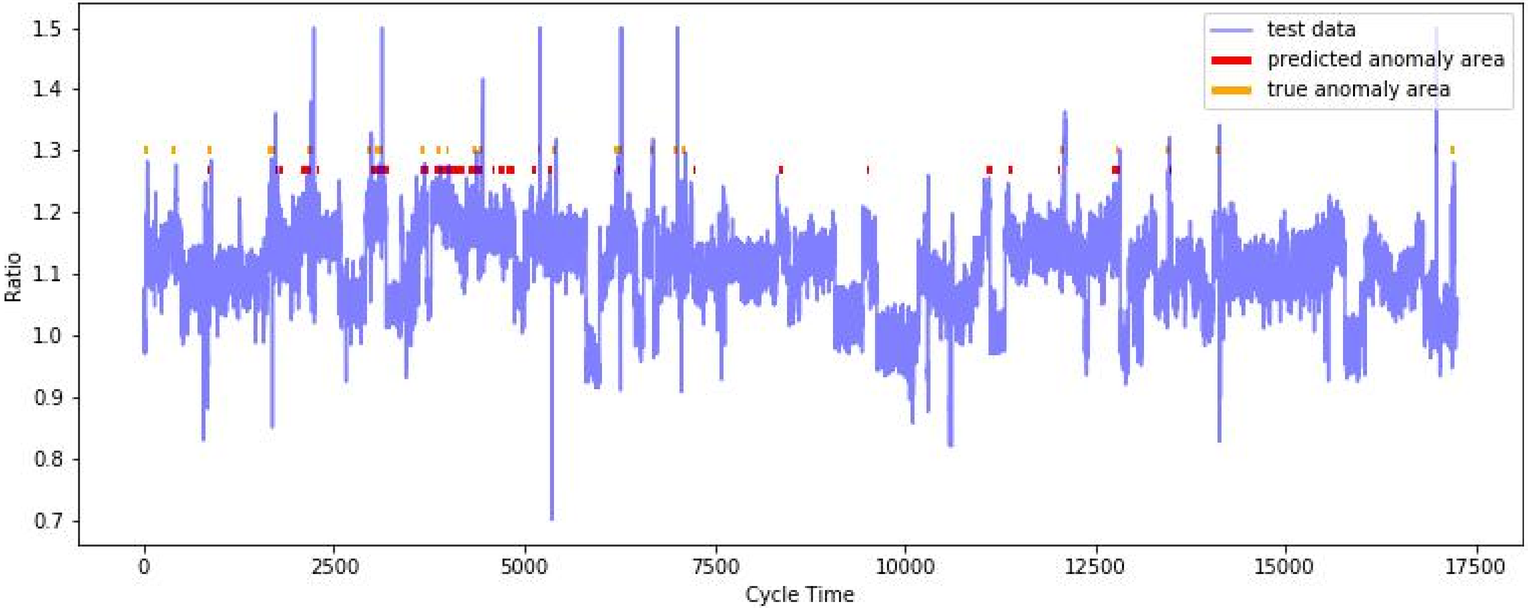
Fitting Result (Zoom in)

- Train set



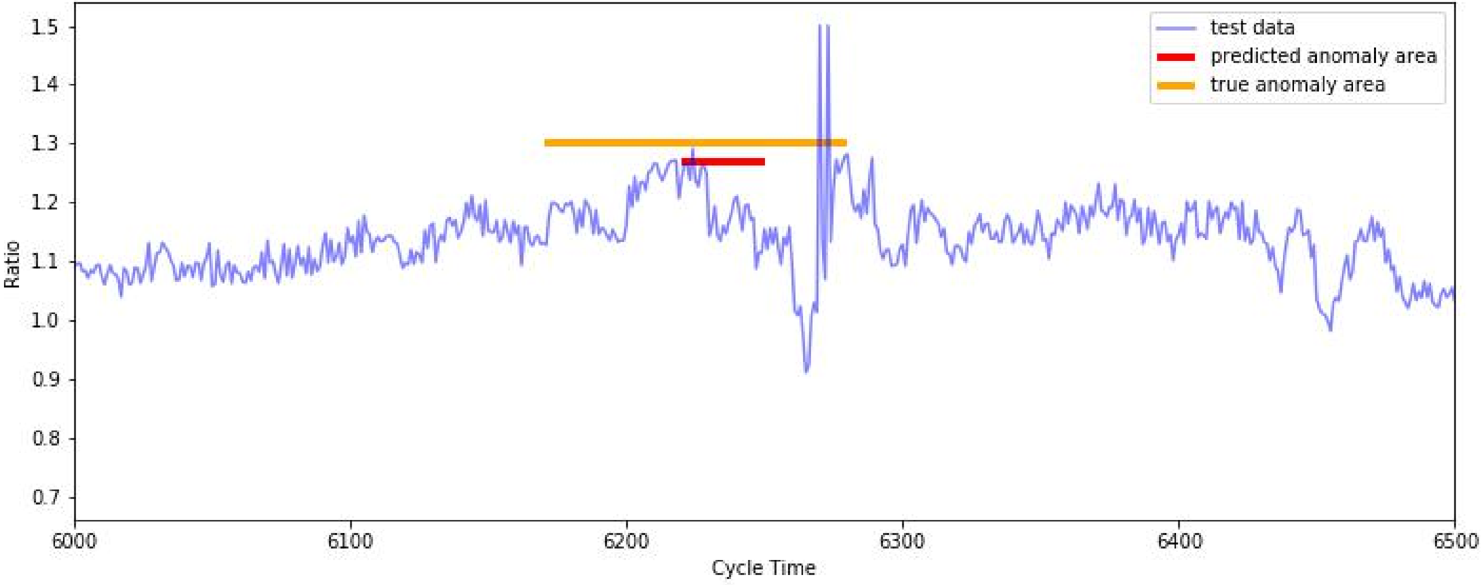
Prediction Result

- Test set



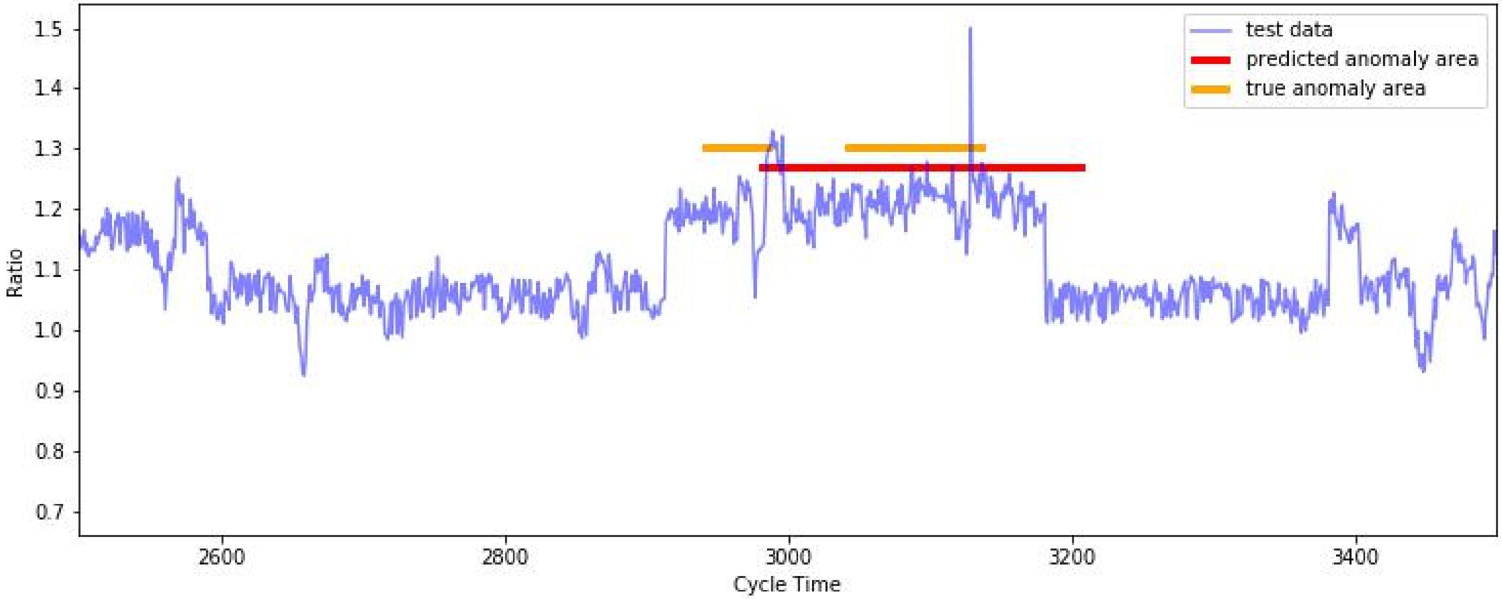
Prediction Result (Zoom in)

- Test set



Prediction Result (Zoom in)

- Test set



Observation

Sand Problem

DOTSON IRON CASTINGS
becoming the world's most agile foundry

Press [Ctrl] +[;] below to add date

Date: 6/7/2019

Shift: 1st Shift

Hour	Production Hours	Taps	Pounds	Molds/Hr	Line 1 1	Line 2 2	Line 3 3	L10 Cycles	L25 Cycles	LFB25 Cycles
6:00 AM				0				44	23	50
7:00 AM				0				74	24	51
8:00 AM				0				85	24	80
9:00 AM				0				64	15	60
10:00 AM				0				138	25	53
11:00 AM				0				150	21	30
12:00 PM				0				112	24	80
1:00 PM				0				101	23	70
2:00 PM				0						
Total	0	0	0	0	0	0	0	768	179	474
Total Target Molds:				Total Cores/Shift:				1421		
				Average Weight Poured:				0.0		
				Average Molds/Hour:				0.0		
				Average Tons/Hour:				0.0		

Core Downtime Summary:

L10:	9 to 9:23 Sand transporter problems
LB25:	8:51 to 9:24 sand transporter problems. Ran Universal C5G three cavity box entire day.
LFB25:	10:45 to 11:18, 11:42 to 11:47 Binder Problems.

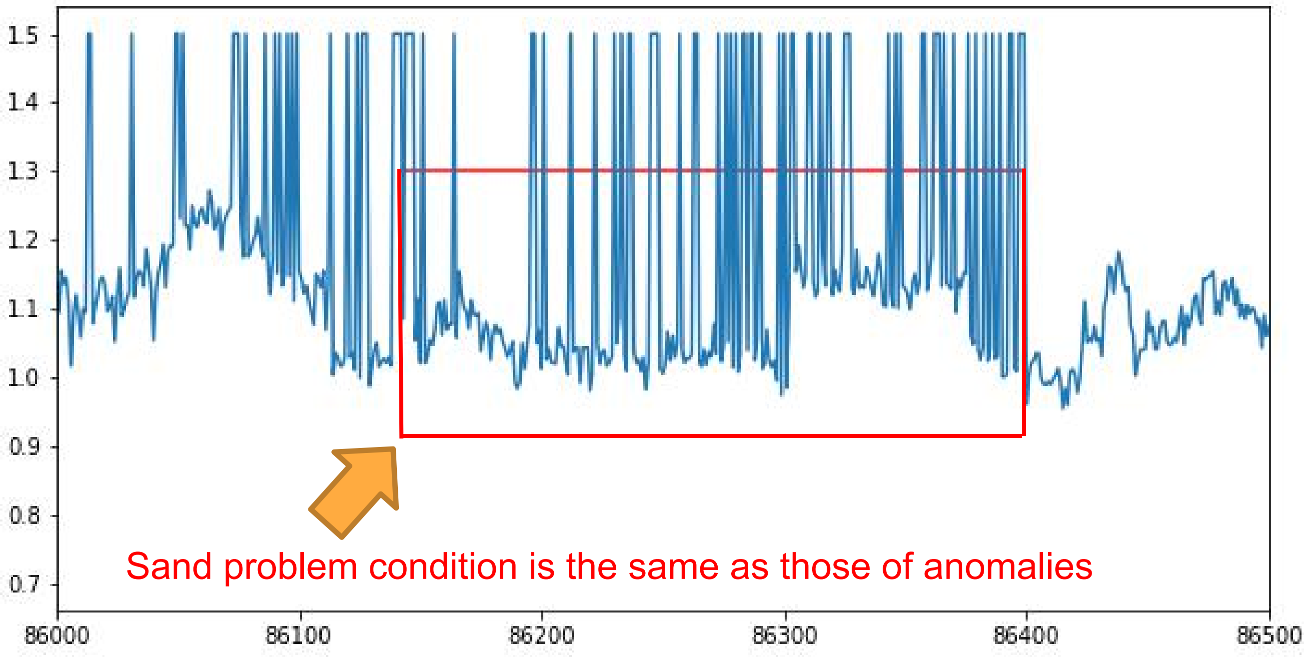
Reason for less than 180 Molds/Hour

LL 10 (core)	LB 25 (core)	LFB 25 (core)	Mold 1	Mold 2	Mold 3	Turntable 1	Turntable 2	Turntable 3	Hoists	Charging	Furnace 1	Furnace 2	Shakeout	Sand Cooler	Hartley	Muller	Sand	Conveyors/Belts	Slow Molding Jobs	Scheduled	Samples	Metal/Machine Mix	Missing People	Breaks

Mold Downtime Summary:

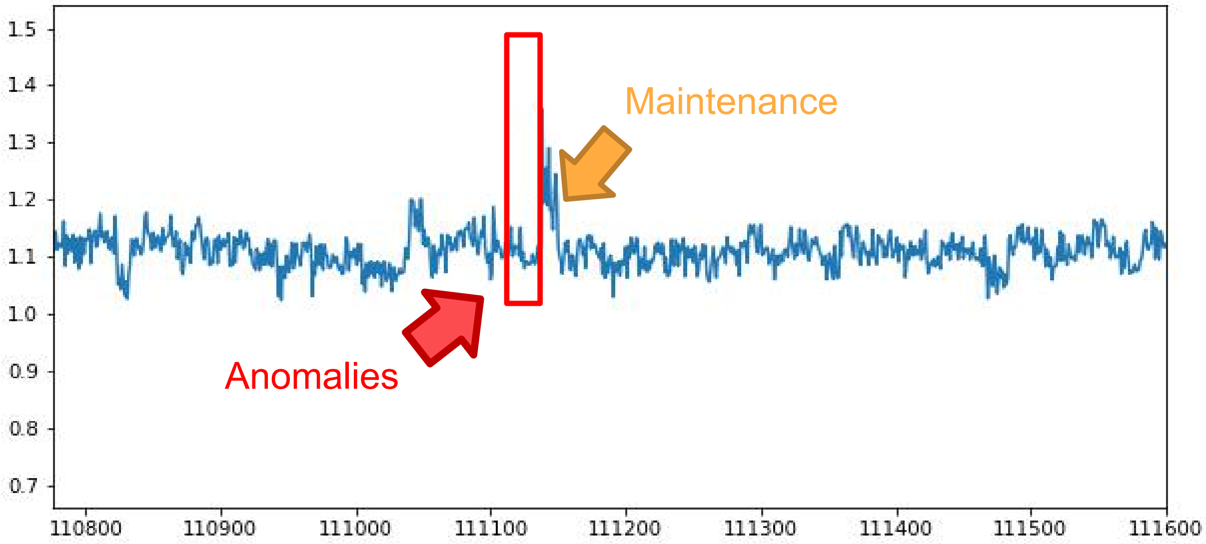
Mold Line 1:	
Mold Line 2:	
Mold Line 3:	

Sand Problem



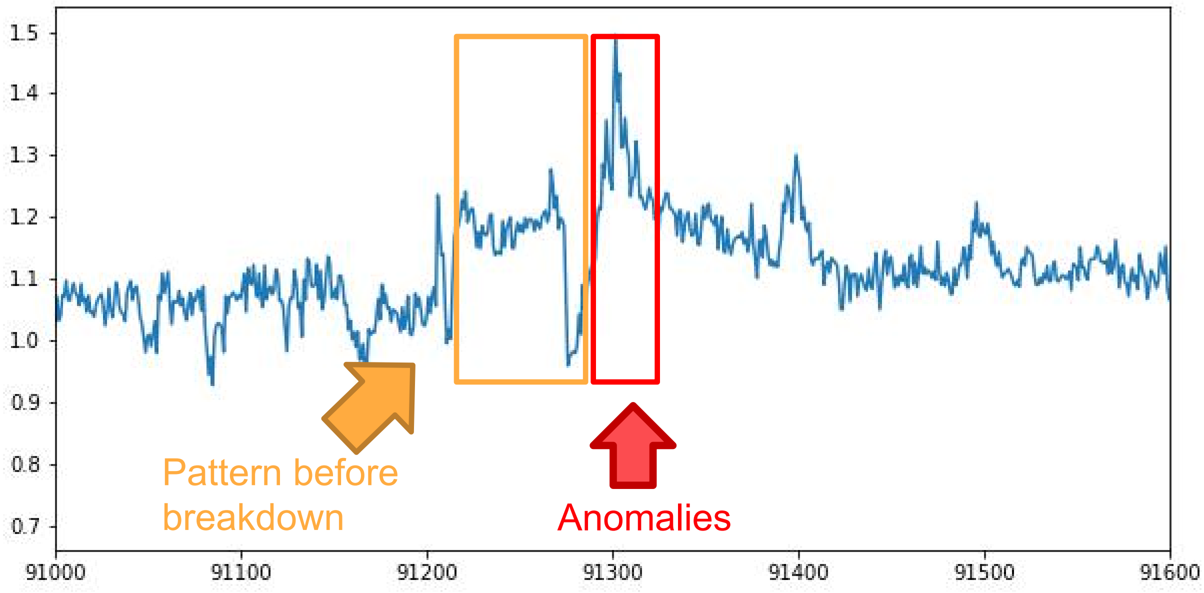
Service and Maintenance

- Existing maintenances in the report are always after the breakdown
- That's too late.



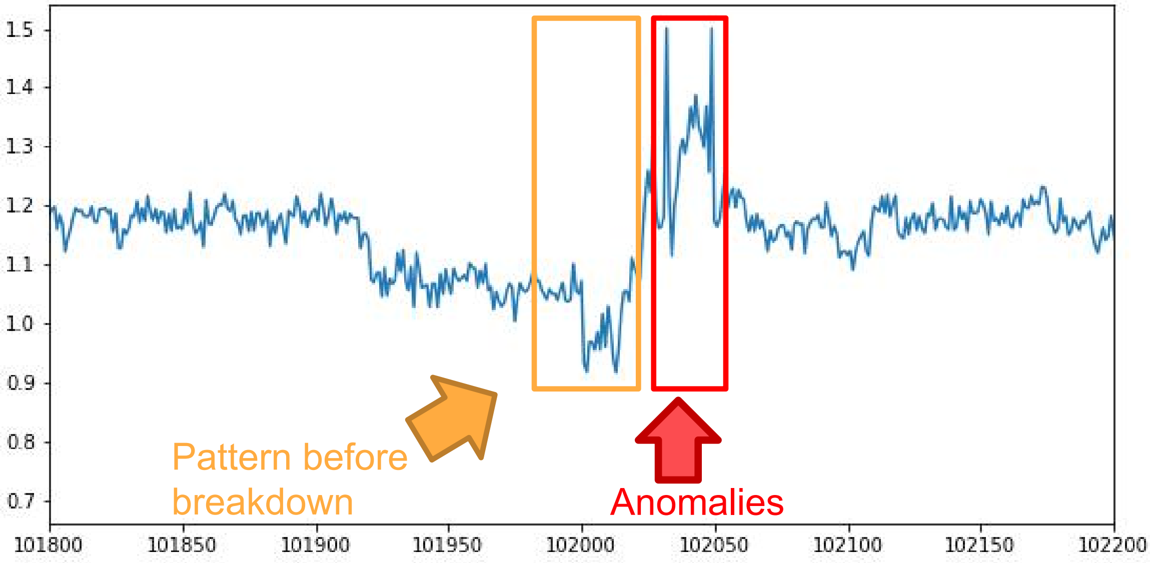
Patterns before breakdown

- 6/17/2019



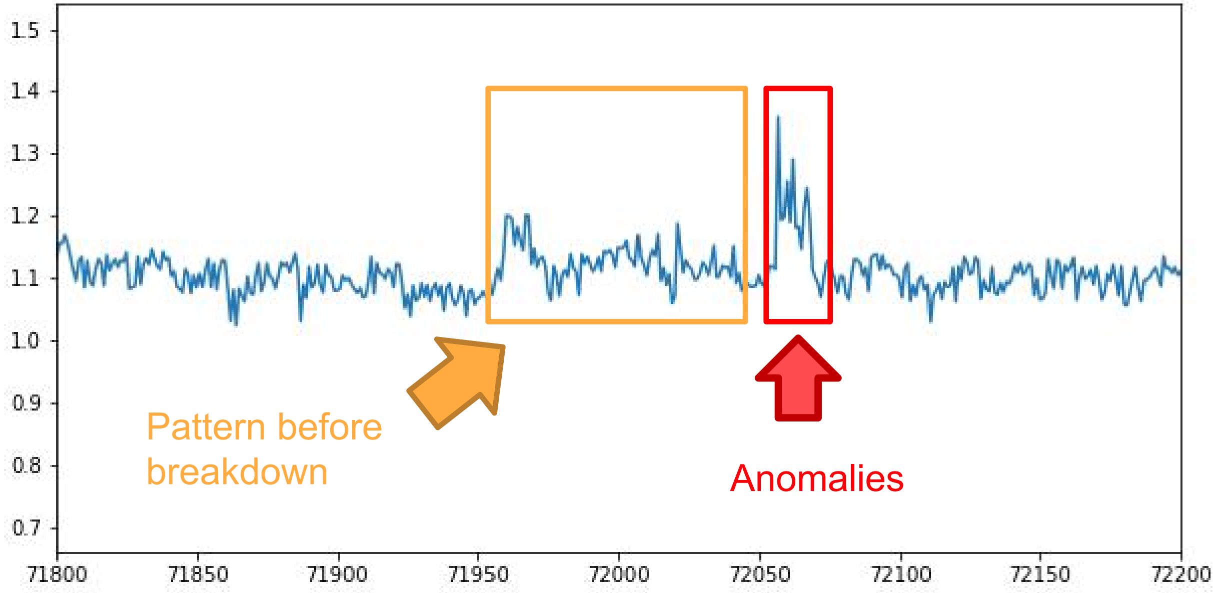
Patterns before breakdown

- 8/5/2019



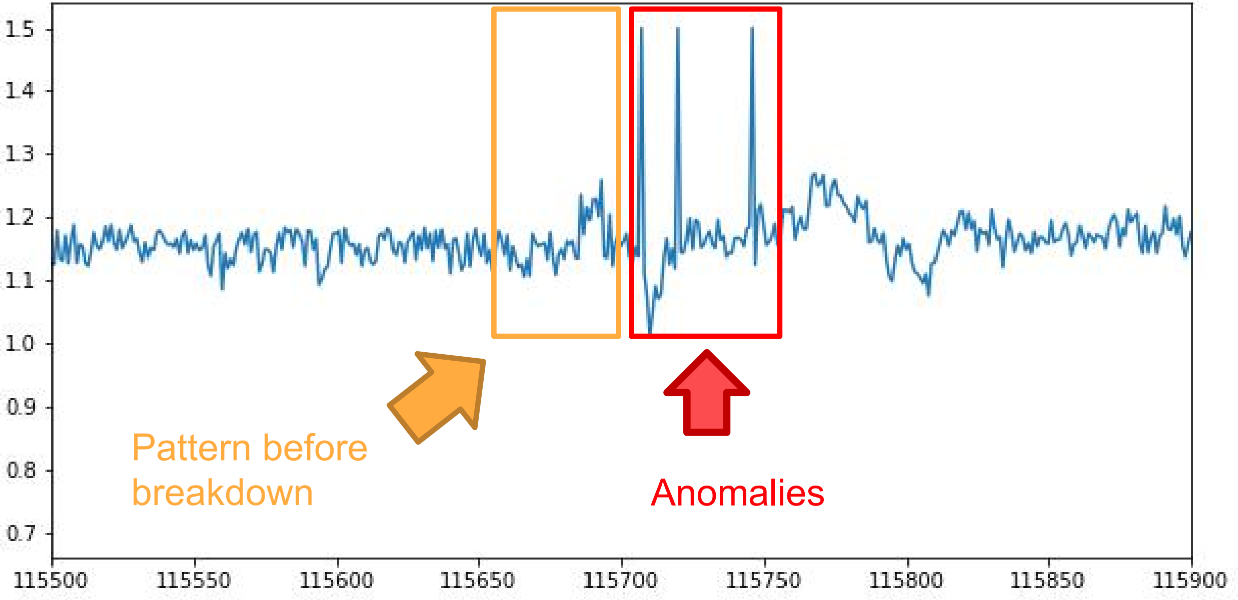
Patterns before breakdown

- 8/20/2019



Patterns before breakdown

- 8/26/2019

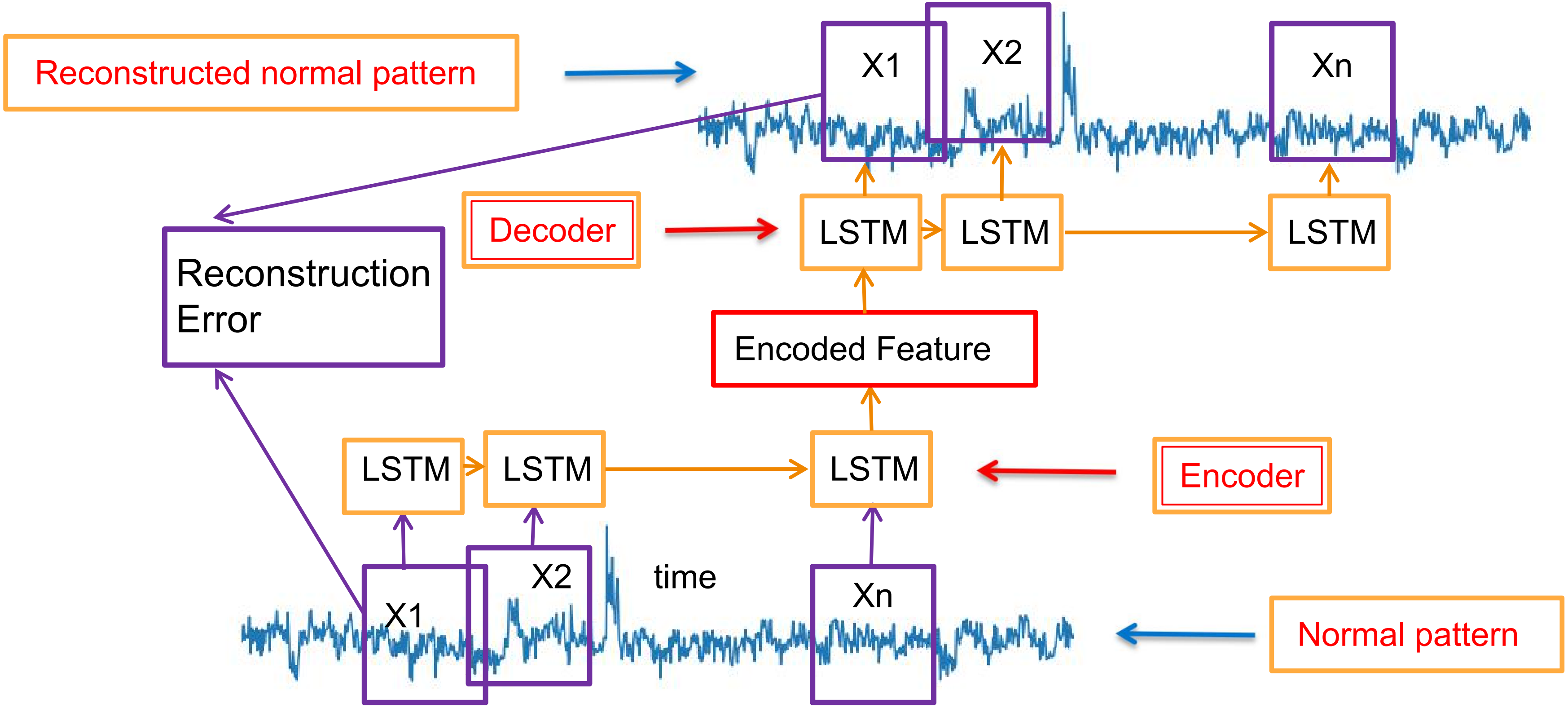


Data

- The data of July only contains that on July 31st.

Deep Learning Method

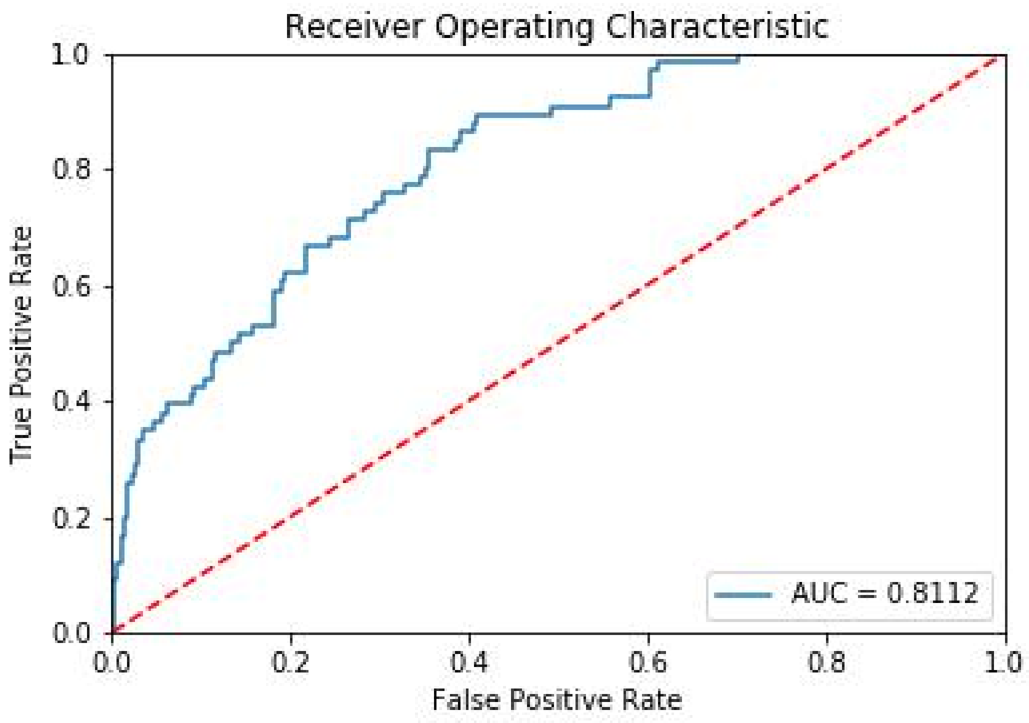
LSTM autoencoder



Choose thresholds

- Use the normal and abnormal data in validation set to determine the thresholds of reconstruction errors.

Preliminary Result



Future Work

Future work

- Model Types
- More metrics to assess the reconstruction error
- Maintenance
- Extreme value theory => test outliers
- Varing the time-steps of the inputs
- Use different Tags together to obtain better results
- ...

- We wish to have more accurate labels for anomalies

Thank you!