UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

CHEMET SCHOOL OF CHEMICAL AND METALLURGICAL ENGINEERING

DANCE Data Analytics, Numerics and Control Engineering Unit

Applications of Data Analytics, Modelling and Control to Solve Industrial Problems. 8 September 2026

Antony Higginson

Kevin Brooks

Contents

- Introduction to DANCE
- Examples
  - Prediction of the effect of Bio-admixtures on concrete workability
  - Non-Linear Dynamic Modelling Of Reagent Effects
  - Optimisation Of Air And Level Profiles In A Flotation Bank
  - Control Of A Flotation Cell At The Point Of Peak Air Recovery
  - Soft Sensors For PGM Grade And Recovery
  - Steam demand forecasting
  - Fault detection in a recycled fibre plant
- Conclusions

Introduction to DANCE

- DANCE stands for Data Analytics, Numeric and Control Engineering
- Around 15 graduate students working mainly in the MMM , paper, petroleum and control fields
- Some 75% of the students are part-time. The majority of the full-time students have some industry funding
- The aim is to use industrial data to solve real-world problems while educating students in both the processes used and the range of tools available

DANCE: Prof Kevin Brooks



- Director at APC Smart
- Has worked at BluEsp, Hatch, Honeywell
- Former President of the South African Council for Automation and Control
- APC application and implementation
- Honorary Adjunct Professor in the School of Chemical and Metallurgical Engineering.

DANCE: Dr Antony Higginson

- Worked in systems integration and process control
- PhD in Robust optimisation of bioethanol production
- Senior Lecturer in the School of Chemical and Metallurgical Engineering
- Teaching Process control and Computing for process Engineers

Prediction of the effect of Bio-admixtures on concrete workability

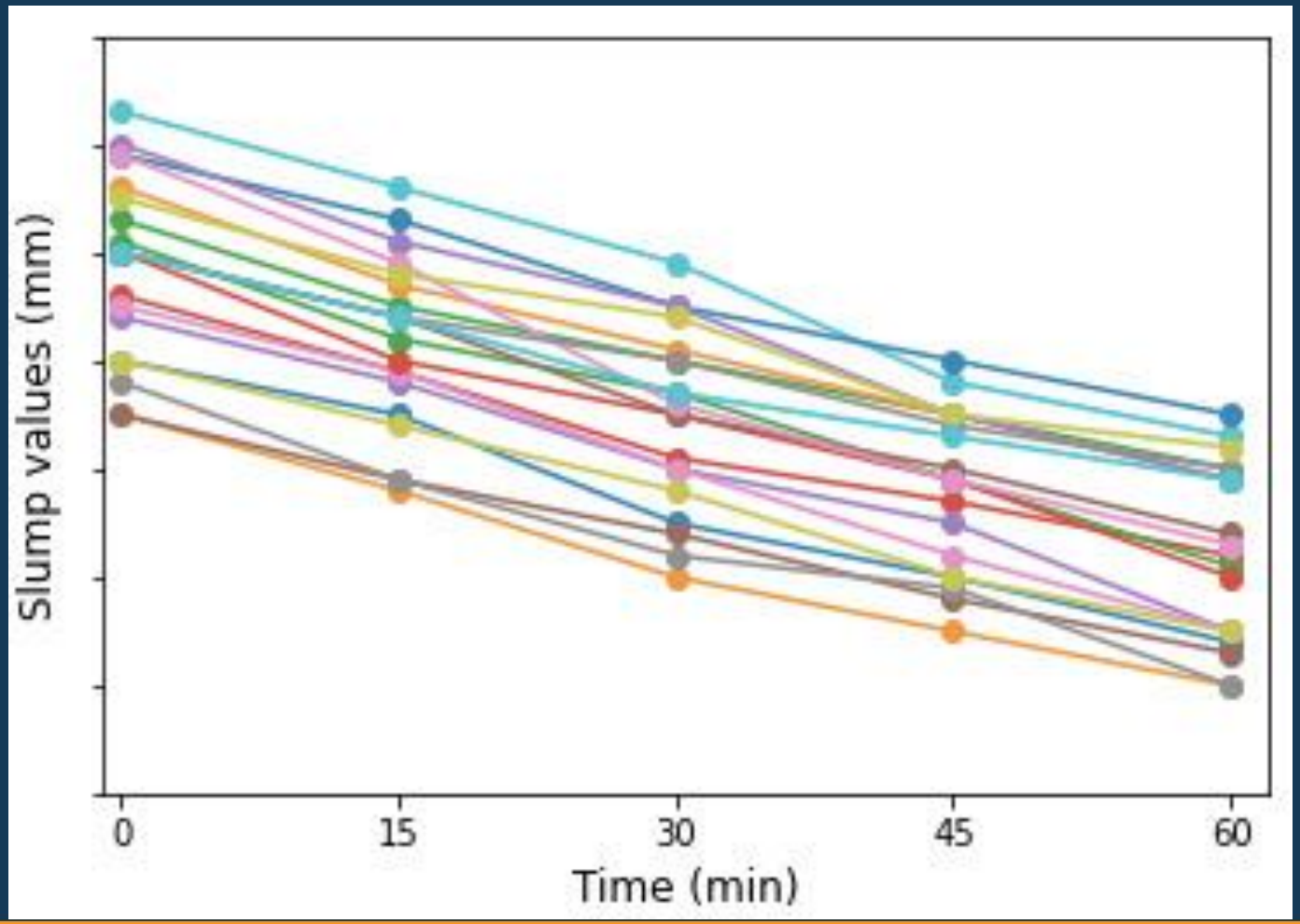
- Problem: A byproduct of wood digestion, Lignosulphonates, can be added to concrete to increase workability. However it is unclear how to predict or measure what the effect the lignosulphonates will be.

Bioadmixtures

- Focused on Slump test and process data
- Slump value should match application



Bioadmixtures



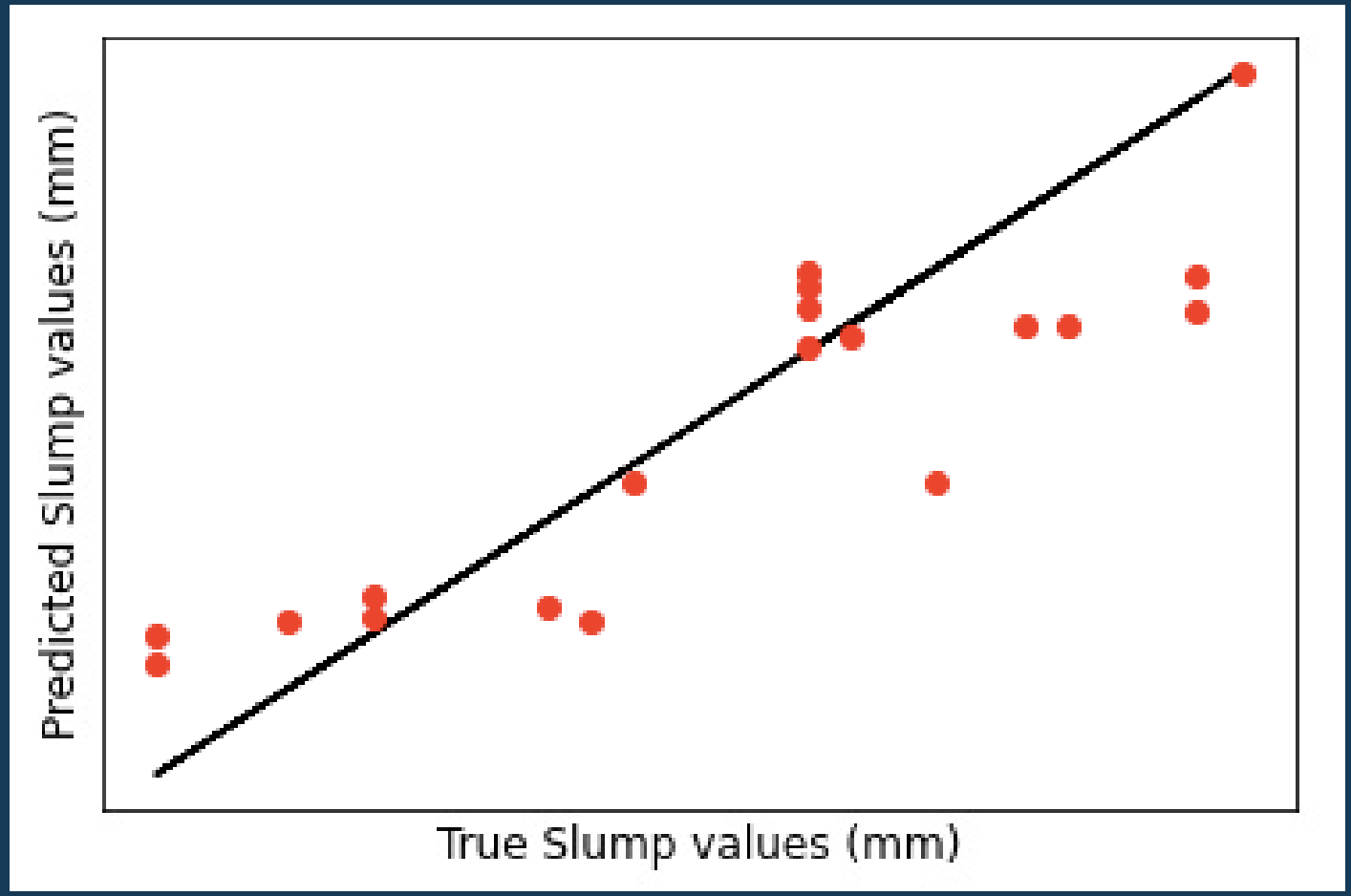
Bioadmixtures

- Linear model for slump with time
- $y = y_o(1 + at)$

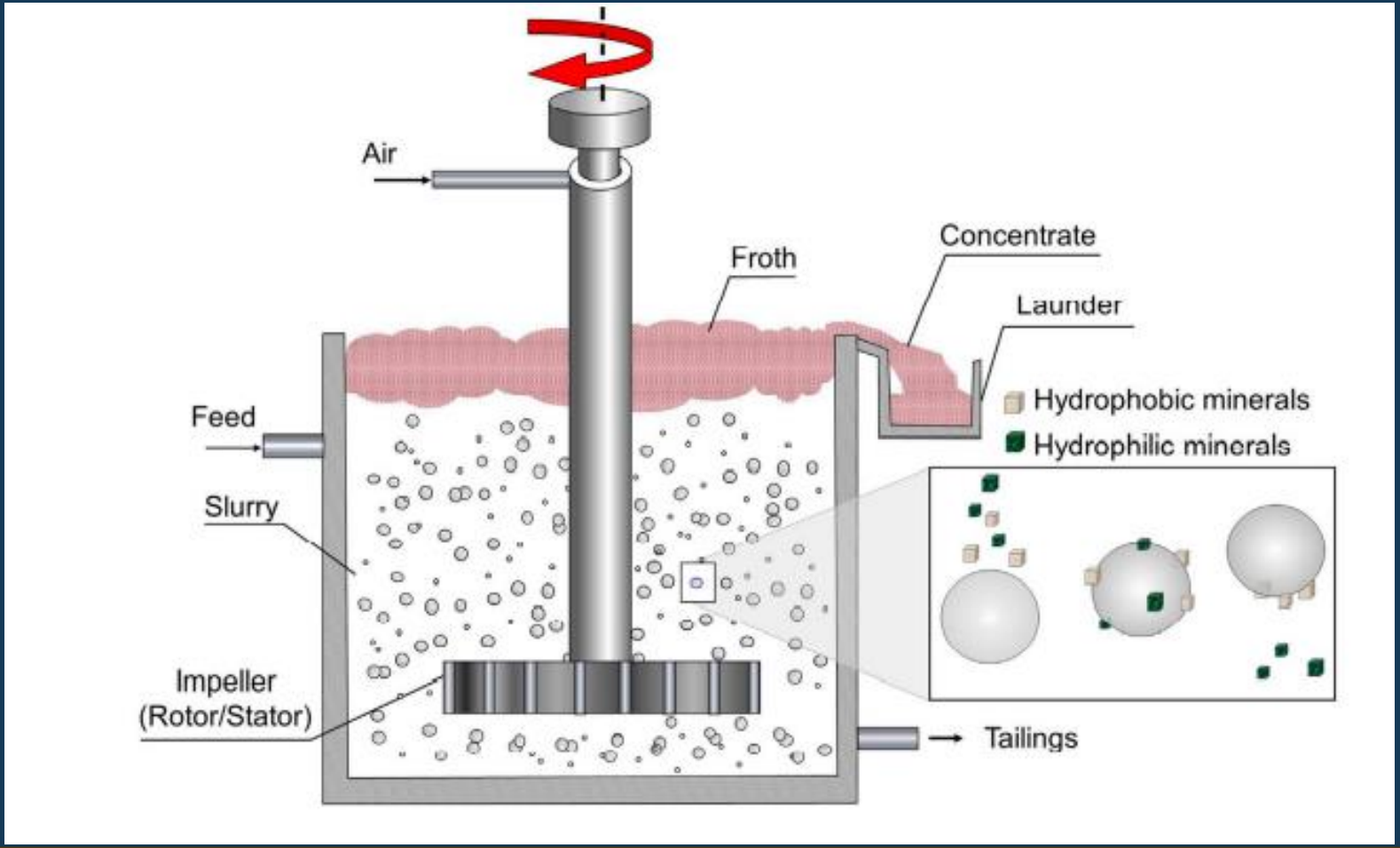
- So we need to be able to predict y_0 based on process data

- Use process data to predict y_0 using SVM

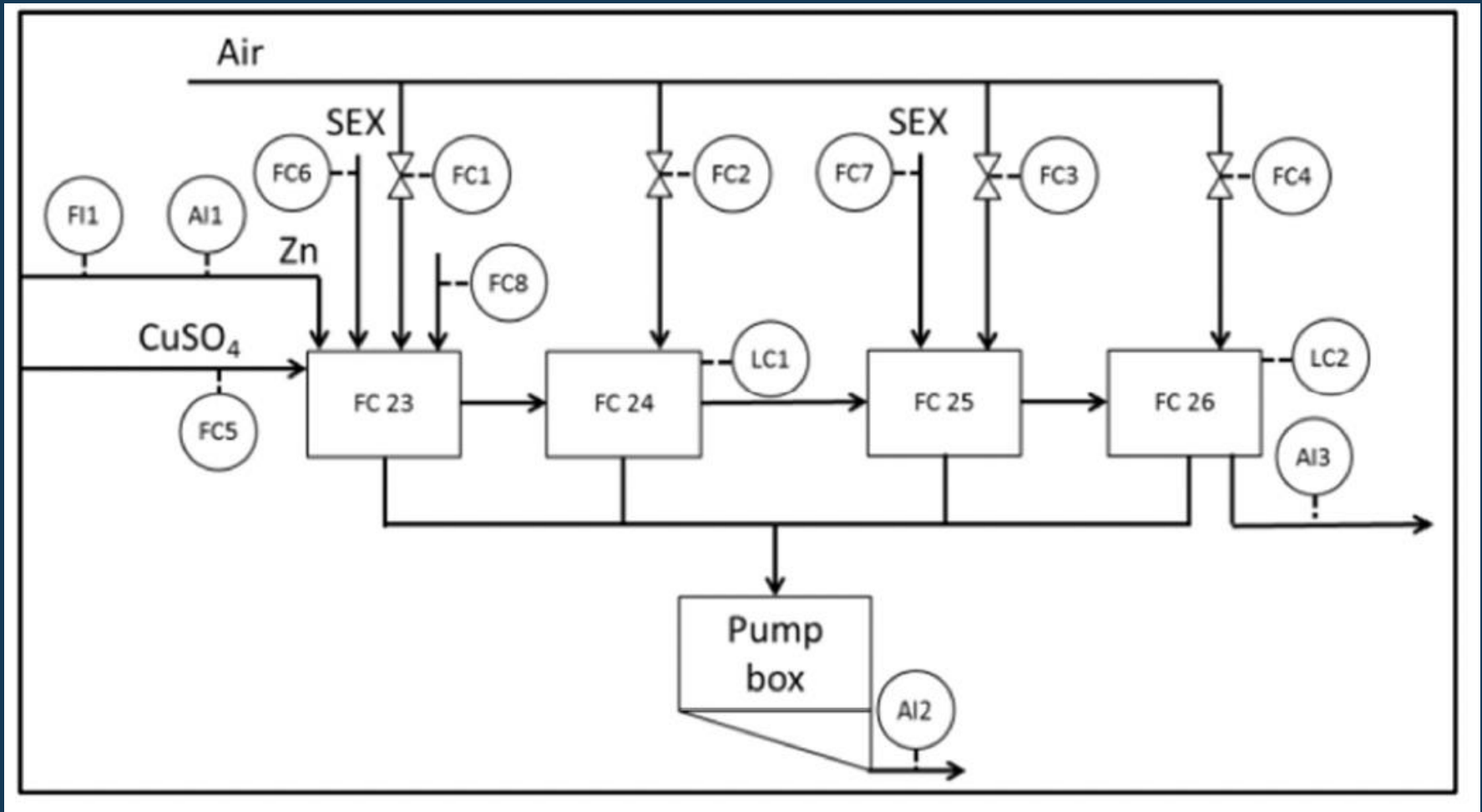
Bioadmixtures



Mineral flotation

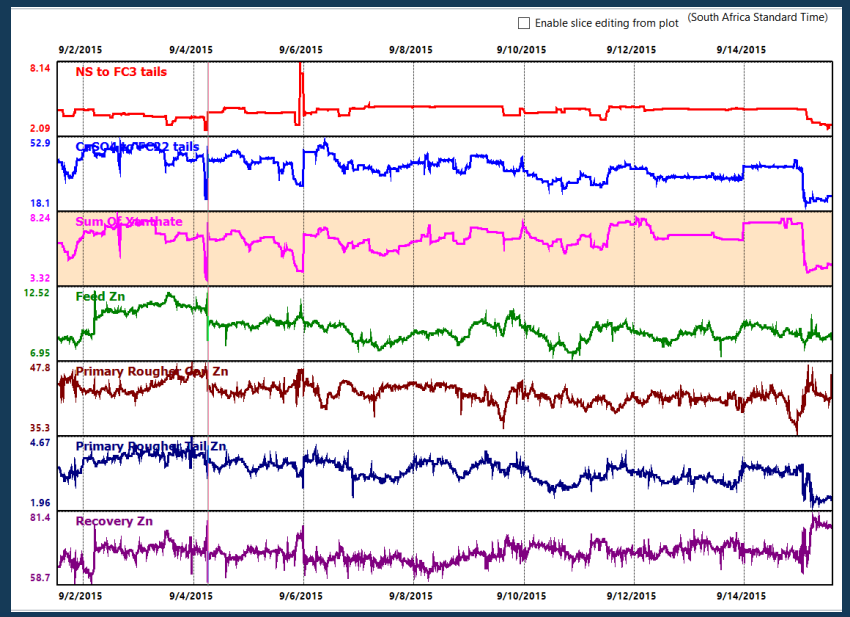


MPC reagent control

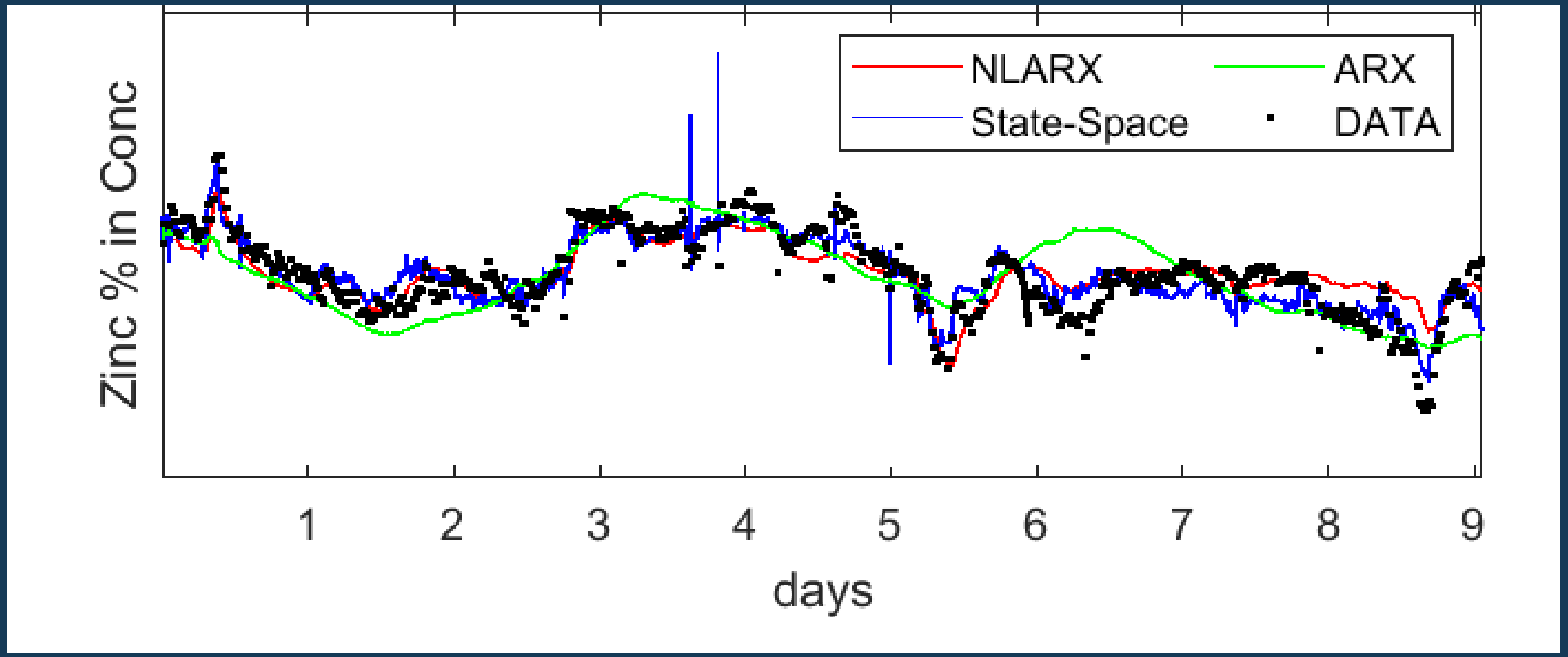


Reagent Effects

- Data from zinc roughers of a lead-zinc plant in Australia (XRF available)
- Reagent step test– air and level held constant
- Linear MPC was implemented, but non-linear responses to reagents observed



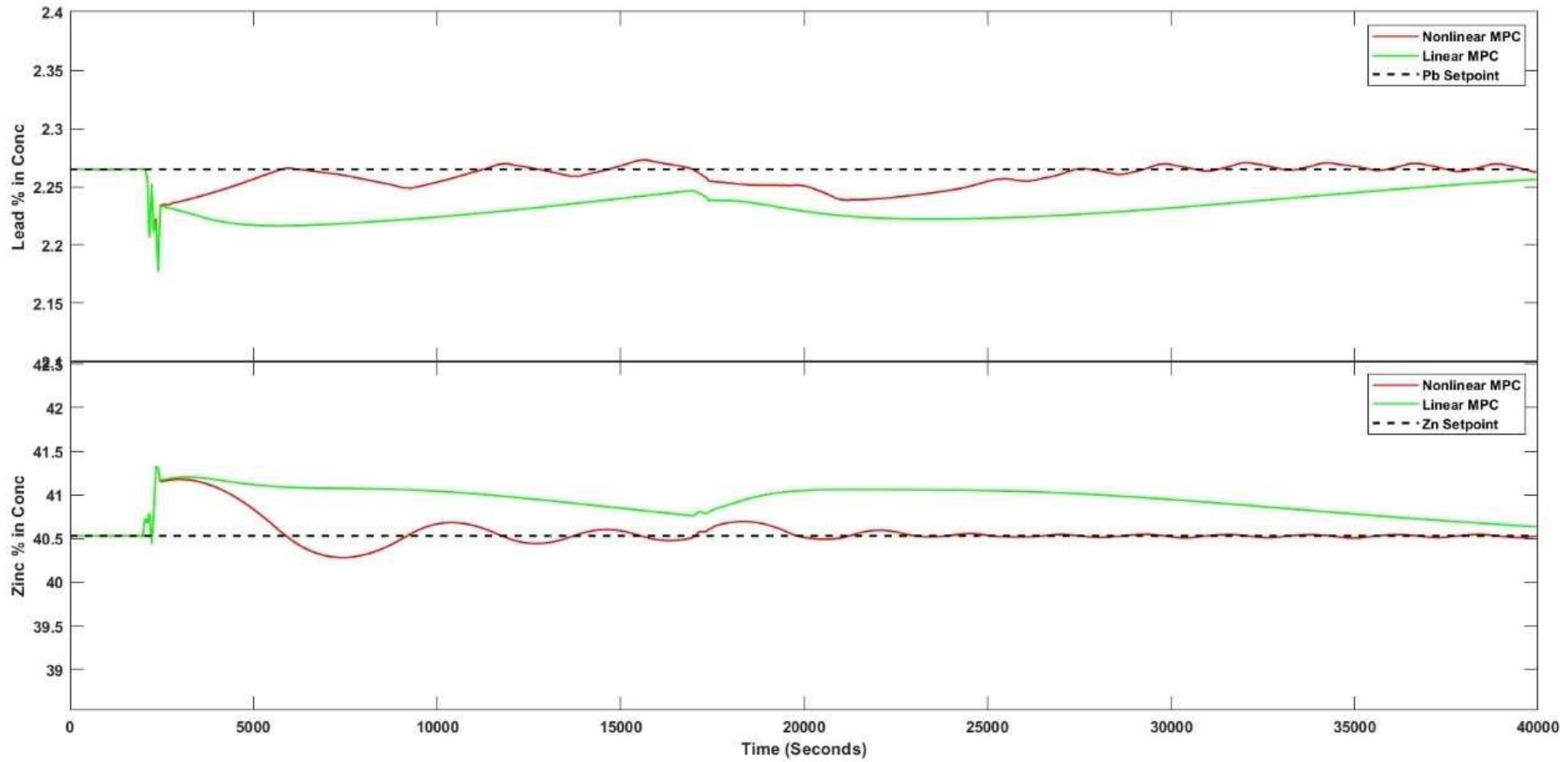
Reagent Effect in Flotation – system identification



Linear vs non-linear models

Variable	Normalised Root Mean Square Error Training Set (%)	Normalised Root Mean Square Error Test Set (%)	R^2 training set	R^2 test set
Con.Zn (SS)	1.9%	6.6%	0.77	0.22
Con.Zn (ARX)	3.3%	12.9%	0.45	0.05
Con.Zn (NLArx)	2.3%	4.4%	0.69	0.38

Implementation of linear and non-linear MPC



Real-Time Dynamic Optimisation Of Air And Level Profiles for a flotation bank

First principles model

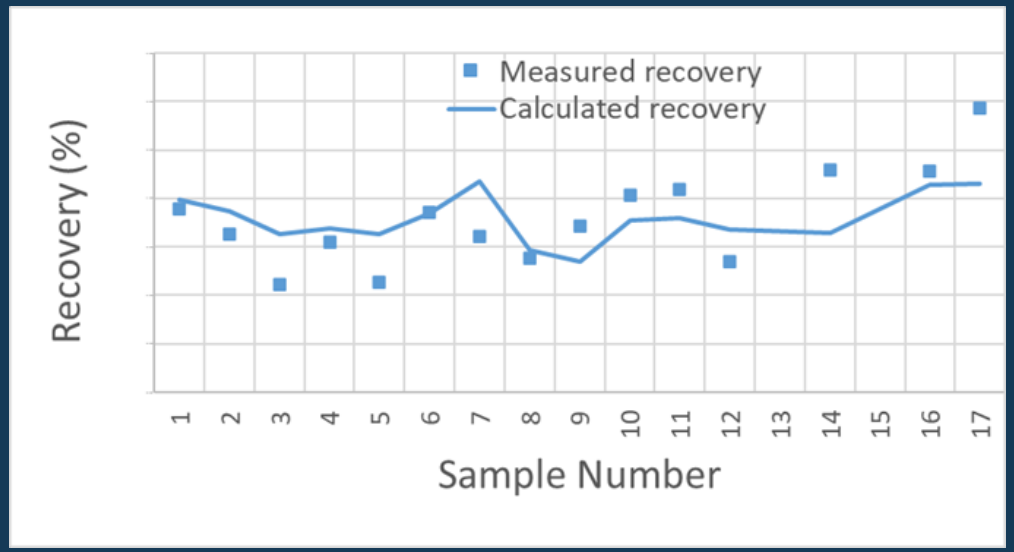
- Mass balance
- Rate expression
- Air recovery
- Hydraulic equations

$$\frac{dm_{pulp,i}}{dt} = m_{in,i} - m_{pulp,i}\left(\frac{F_{in}}{V_{pulp}} + k_i\right)$$

$$k = \frac{6J}{d}\alpha AR$$

$$AR = A' + B'\frac{h}{J} + C'J$$

Fit for nine cells in a rougher bank

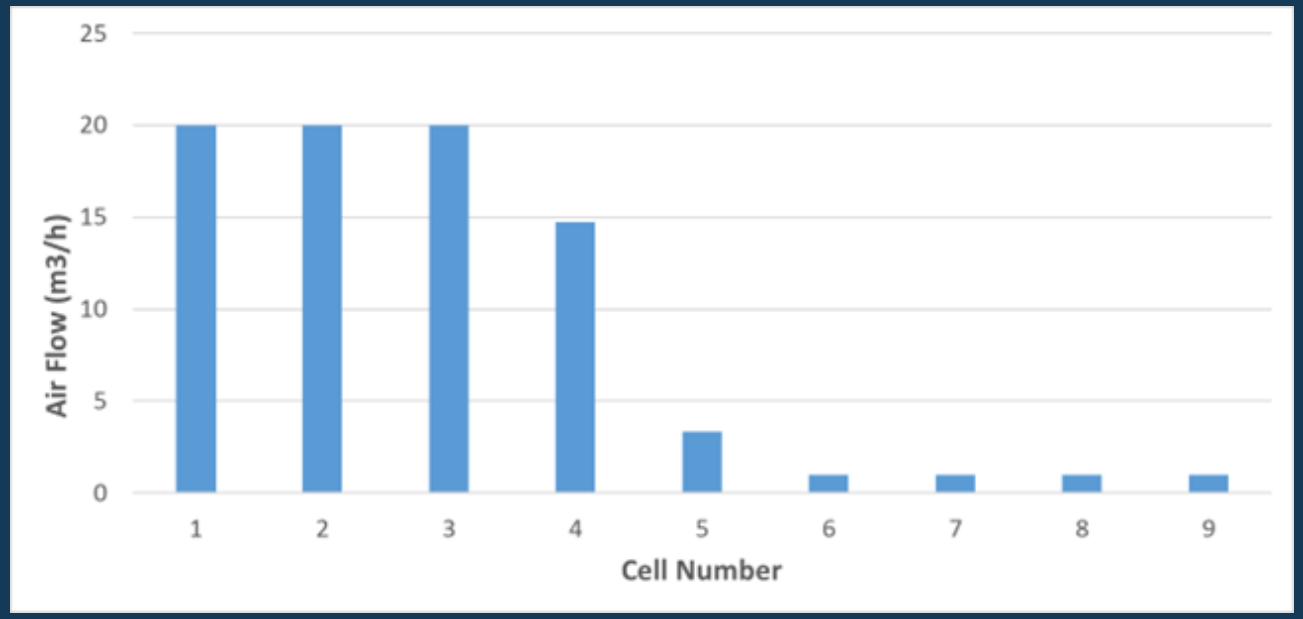


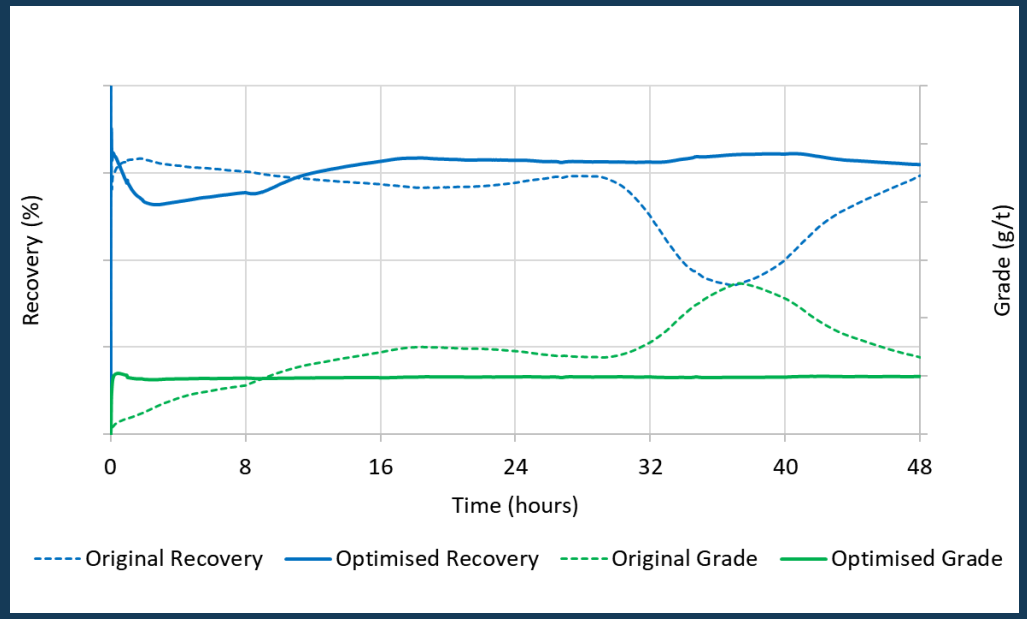
Metric	Training Set
Root Mean Square Error	1.6
R^2	0.29

- Fits are not very good, possibly because of measurement errors and the fact that no reagent effects have been included
- Parameters give qualitatively correct behaviour and are therefore used in simulation

Optimisation Results

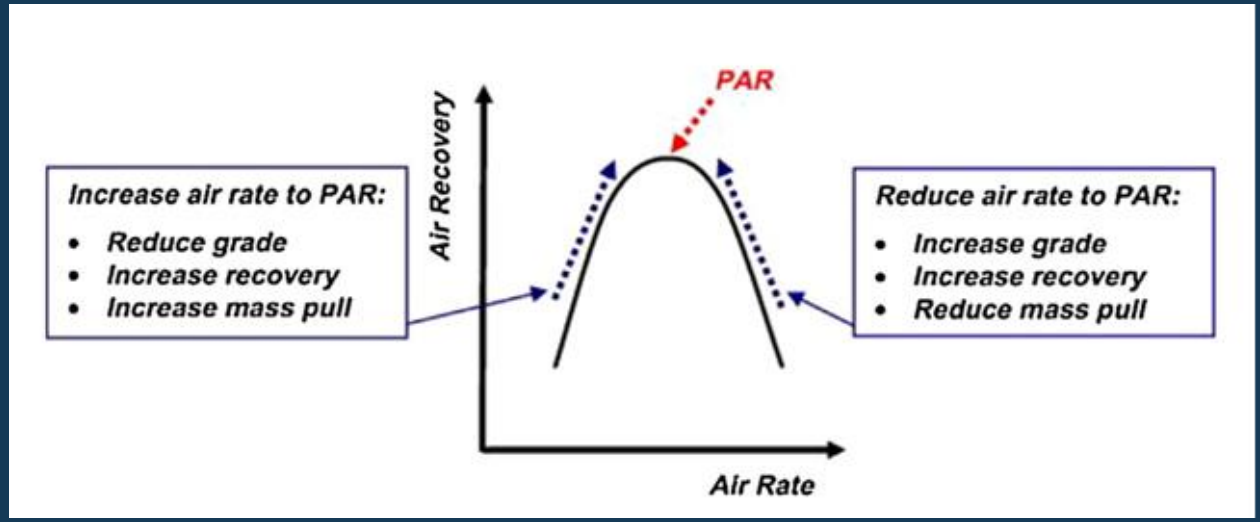
- Recovery maximised subject to minimum grade constraint, minimum and maximum air flows





Control Of A Flotation Cell At Peak Air Recovery

- Air recovery is the fraction of air that leaves the pulp phase of a flotation cell and reports to the concentrate
- Work from the flotation group at Imperial College indicated that recovery could be maximised by operating at the *peak air recovery(PAR)* point



A Possible Solution

- If we had a model for air recovery, maybe we could find the peak
- Such a model has been developed by Oosthuizen et al. (2021). The steady-state form of the model looks like:
- $\alpha = A + B\frac{h}{J} + CJ$

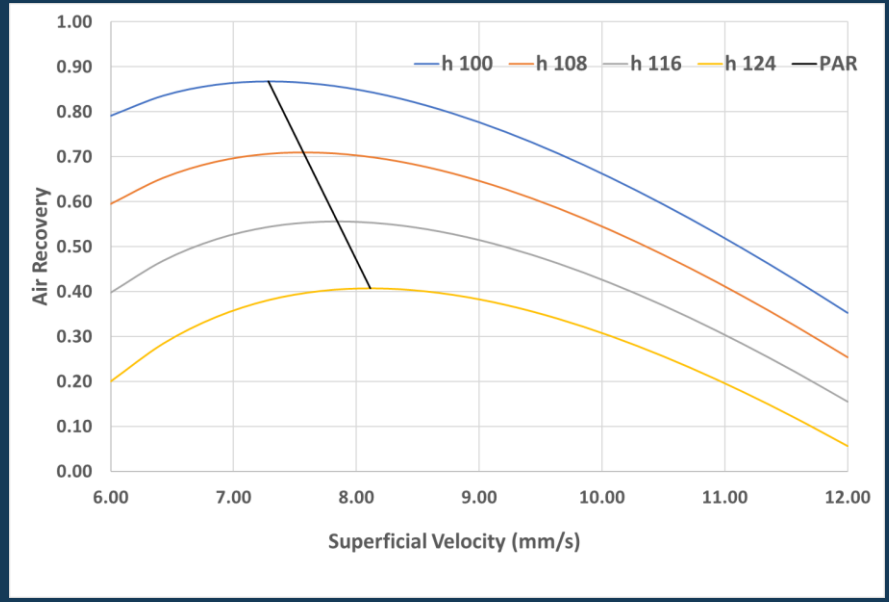
were

α = air recovery

h= froth depth (m)

J= superficial gas velocity (m/s)

A, B, C are empirical constants



Control Scheme

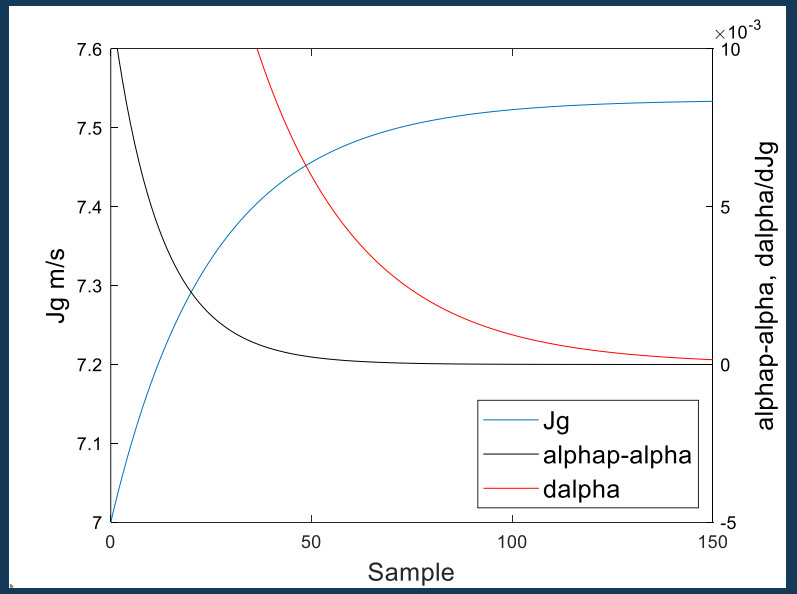
- α can be measured on a plant with a camera that measures the froth velocity and a laser that measures the froth height
- So we take a number of measurements (say 10) and use them to find the coefficients A, B and C by regression
- Now we find the derivative of our equation wrt J:
- $\frac{d\alpha}{dJ} = \mathrm{C} - \frac{B}{J^2}$
- We use $d\alpha_P/dJ$ as the process variable for a PID loop, with setpoint zero, that manipulates J until we reach the peak

Matrix based regression

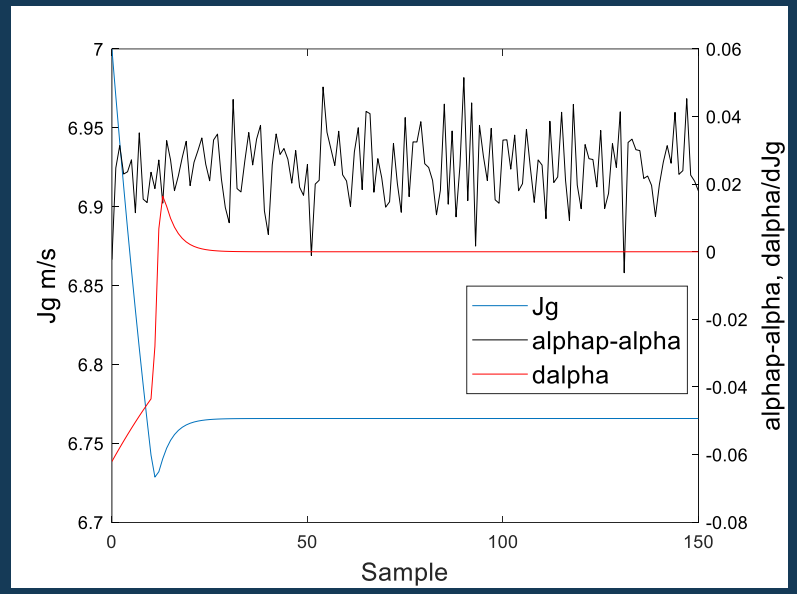
- α can be measured on a plant with a camera that measures the froth velocity and a laser that measures the froth height
- So we take a number of measurements (say 10) and use them to find the coefficients A, B and C by regression
- Now we find the derivative of our equation wrt J:
- $\frac{d\alpha}{dJ} = \mathrm{C} - \frac{B}{J^2}$
- We use $d\alpha_P/dJ$ as the process variable for a PID loop, with setpoint zero, that manipulates J until we reach the peak

Results

- A simulation was developed in Matlab Simulink



Results with no noise



Results with some noise

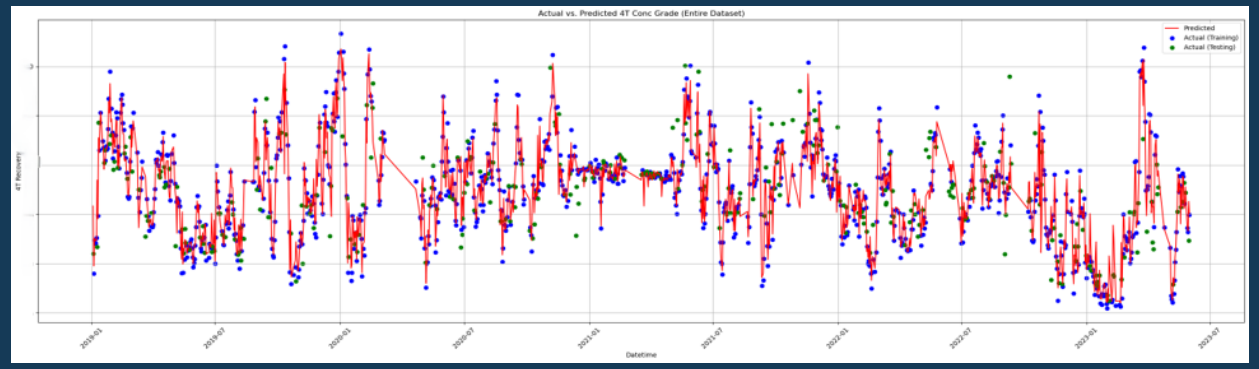
Soft Sensors For PGM Grade And Recovery

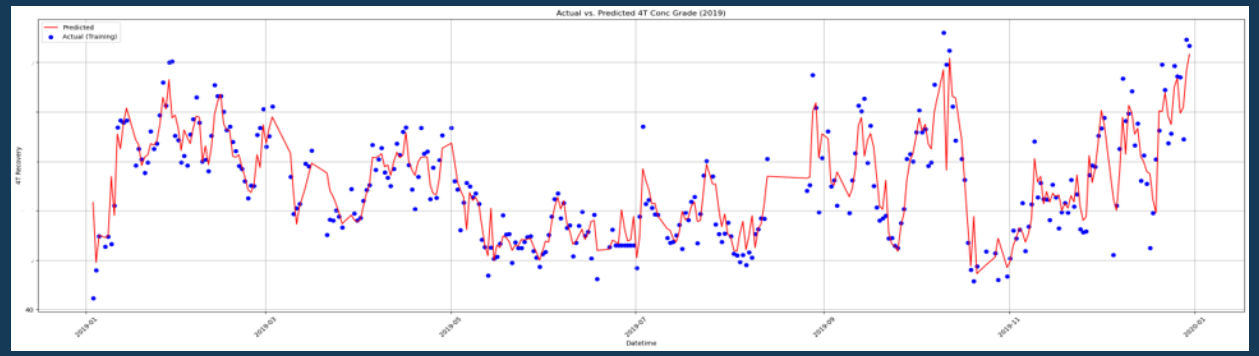
- A challenge to the application of advanced control techniques is the lack of high-frequency measurements of grade and recovery.
- Grade measurements may be only once per shift and are typically composite samples.
- This study investigated supervised learning algorithms that can provide soft sensors for grade and recovery on a rougher bank of a platinum group metals (PGM)) concentrator

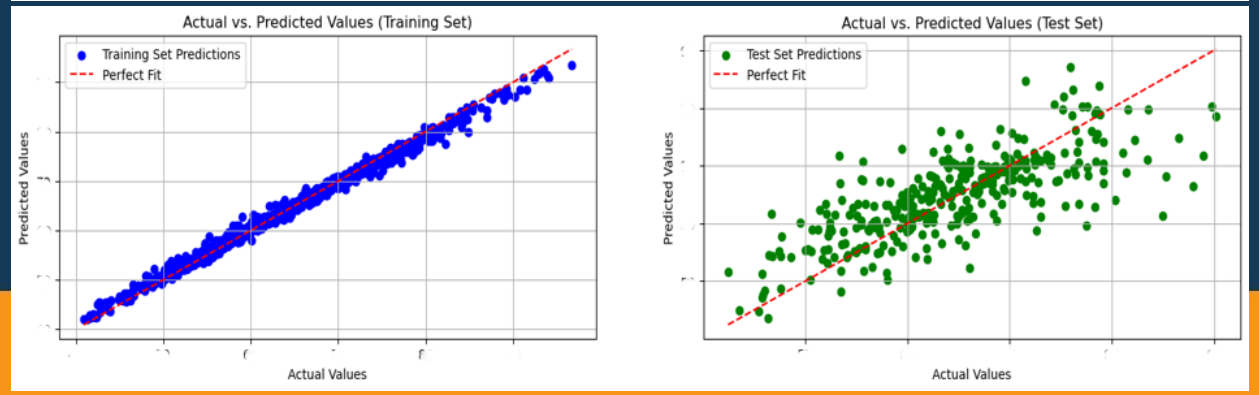
Machine learning approach

- Data for four years of operation of a rougher bank of seven cells was gathered
- After data smoothing, a feature selection algorithm was used to determine the most important parameters to use as inputs to the model.
- The algorithm selects a combination of process variables, including feed and air flows, reagent flows, and some derived measurements from the froth cameras.
- A total of 16 inputs were selected.

Results – Recovery (XGBoost model)

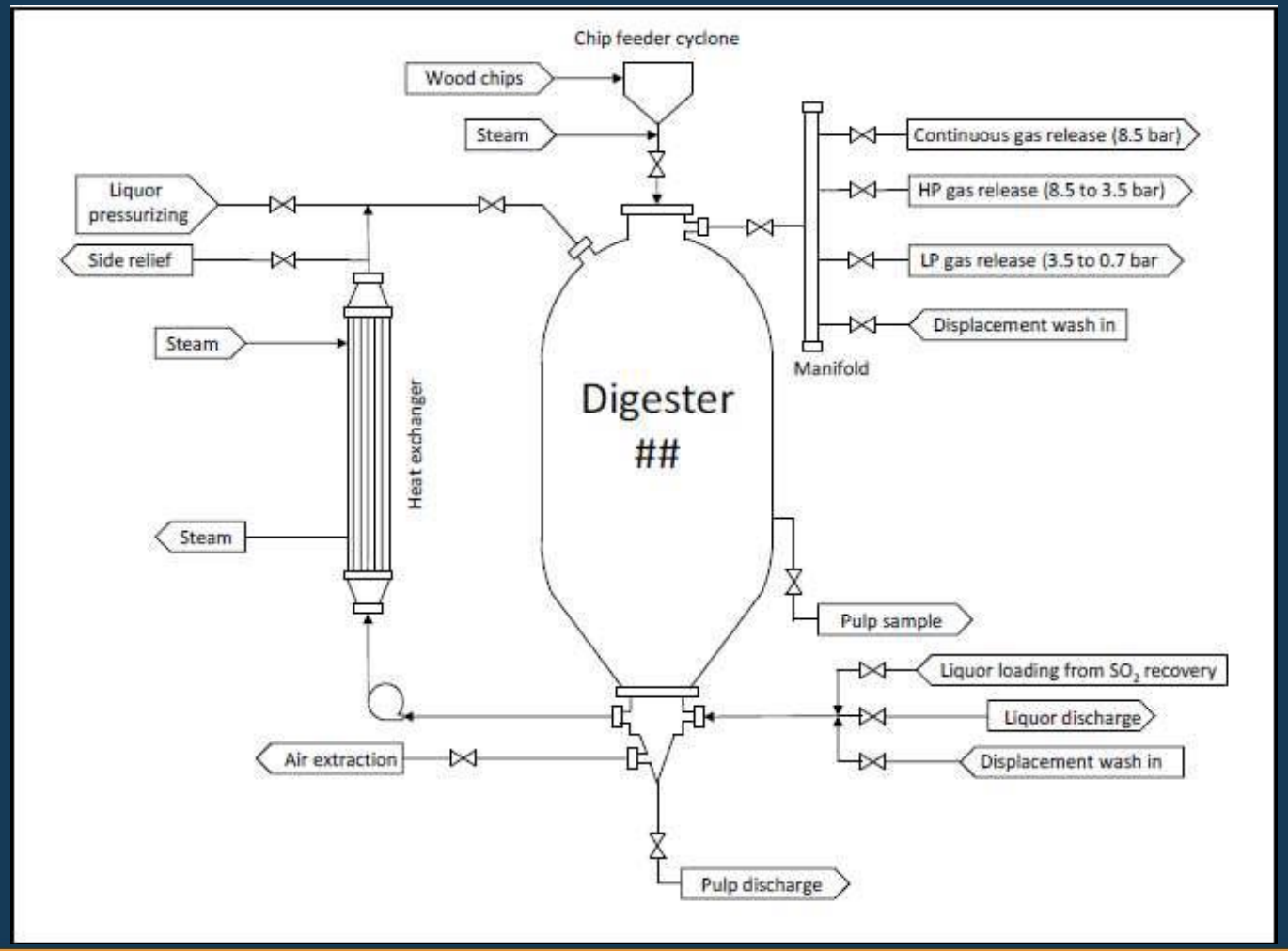




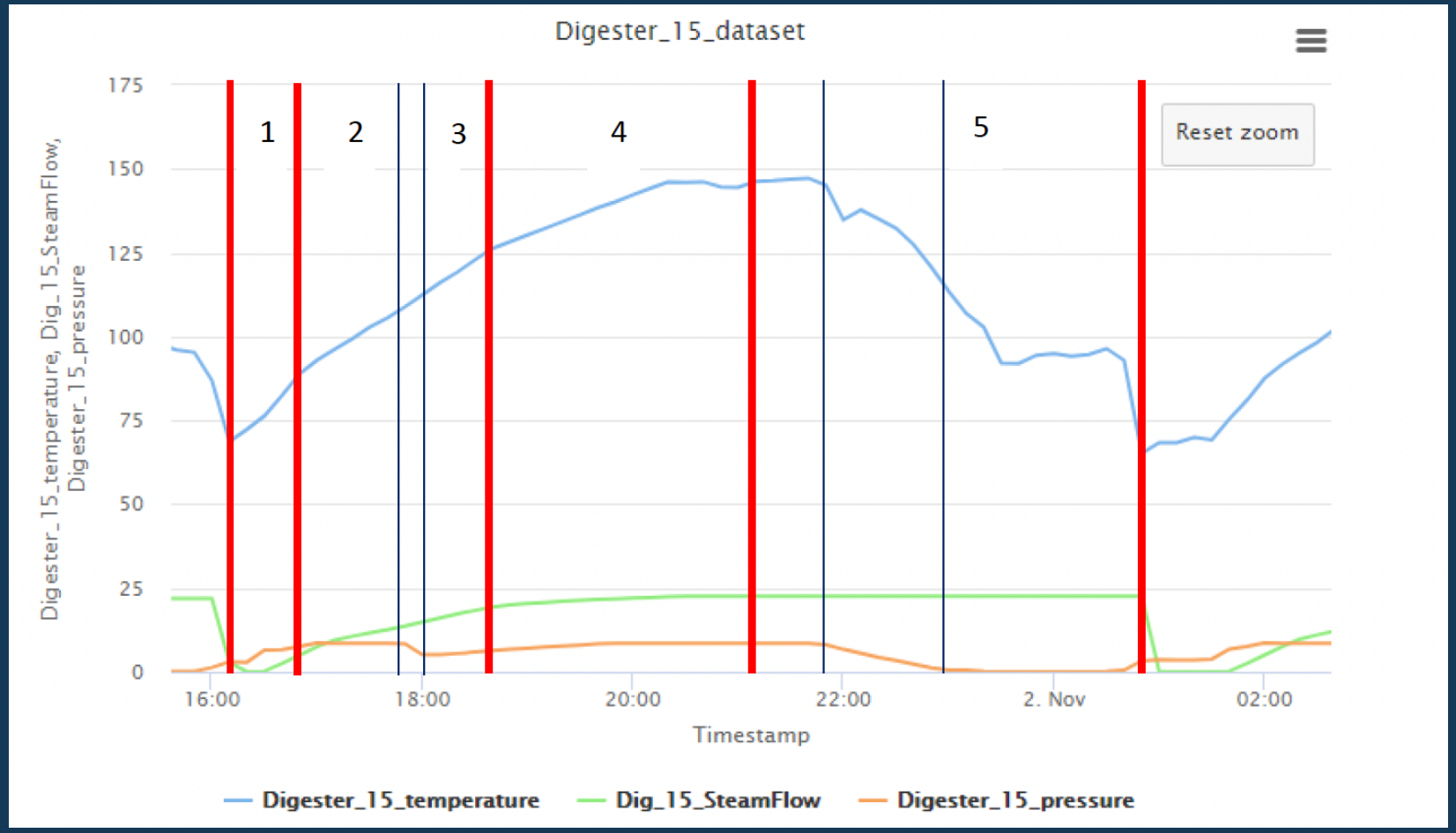


Evaluation Metrics XGBoost Regression		
	Training Set	**Testing Set**
RMSE	1.03	6.05
R^2	0.99	0.56

Steam demand Forecasting at SAICorr



Stem demand forecasting



Modelling

$$\frac{dT}{dt} = \frac{\dot{Q} + (-\Delta H_{rx})(N_{w0} k e^{-kt})}{\sum N_i C_{Pi}}$$

$$\frac{dT}{dt} = \frac{\dot{Q} + (-\Delta H_{rx})(N_{w0} k e^{-kt})}{N_{W0}(C_{P0} + y_{W0} X (C_{PP} - C_{PW}))}$$

Let $\theta = \frac{T - T_0}{T_0}$ and $\tau = \frac{t}{t_b}$

$T = \theta T_0 + T_0$ and $t = \tau t_b$

$dT = T_0 d\theta$ and $dt = t_b d\tau$

$$\frac{d\theta}{d\tau} = \frac{\sigma}{1 + y_{W0} X \gamma} + \frac{\beta (k t_b e^{-k\tau t_b})}{1 + y_{W0} X \gamma}$$

$kt_b = \boldsymbol{Da}$

$$\frac{d\theta}{d\tau} = \frac{\sigma}{1 + y_{W0} X \gamma} + \frac{\beta (Da e^{-Da\tau})}{1 + y_{W0} X \gamma}$$

Steam demand forecasting

Let $\boldsymbol{e^{-Da\tau}} = \boldsymbol{\rho}$

$$\frac{d\theta}{d\tau} = \frac{\sigma + \beta Da\rho}{1 + y_{W0}X\gamma}$$

Simplifying further:

Let $\beta Da = \boldsymbol{\delta}$

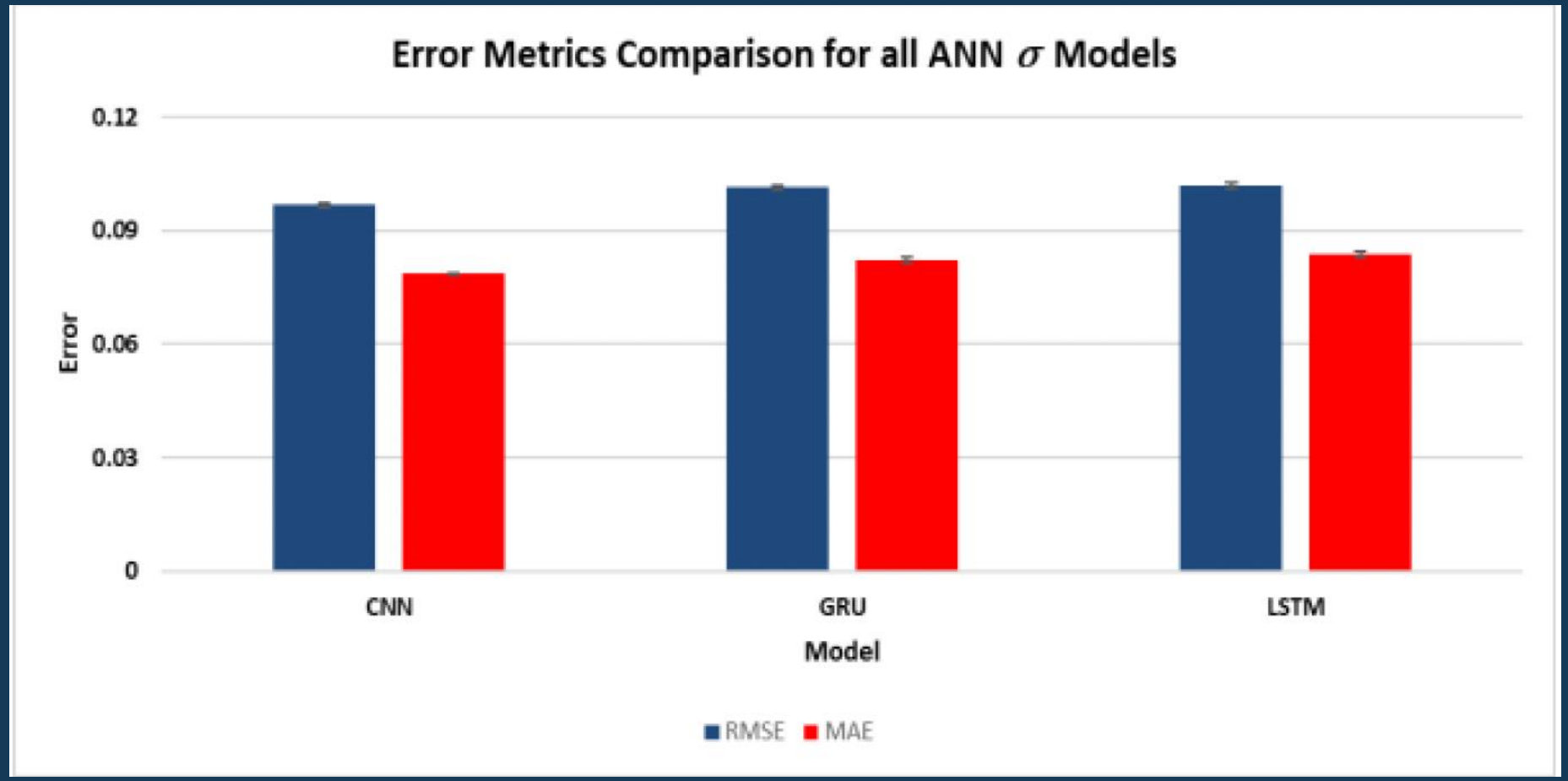
$y_{w0}\gamma = \boldsymbol{\alpha}$

$X = 0.47$ (Sappi, 2020)

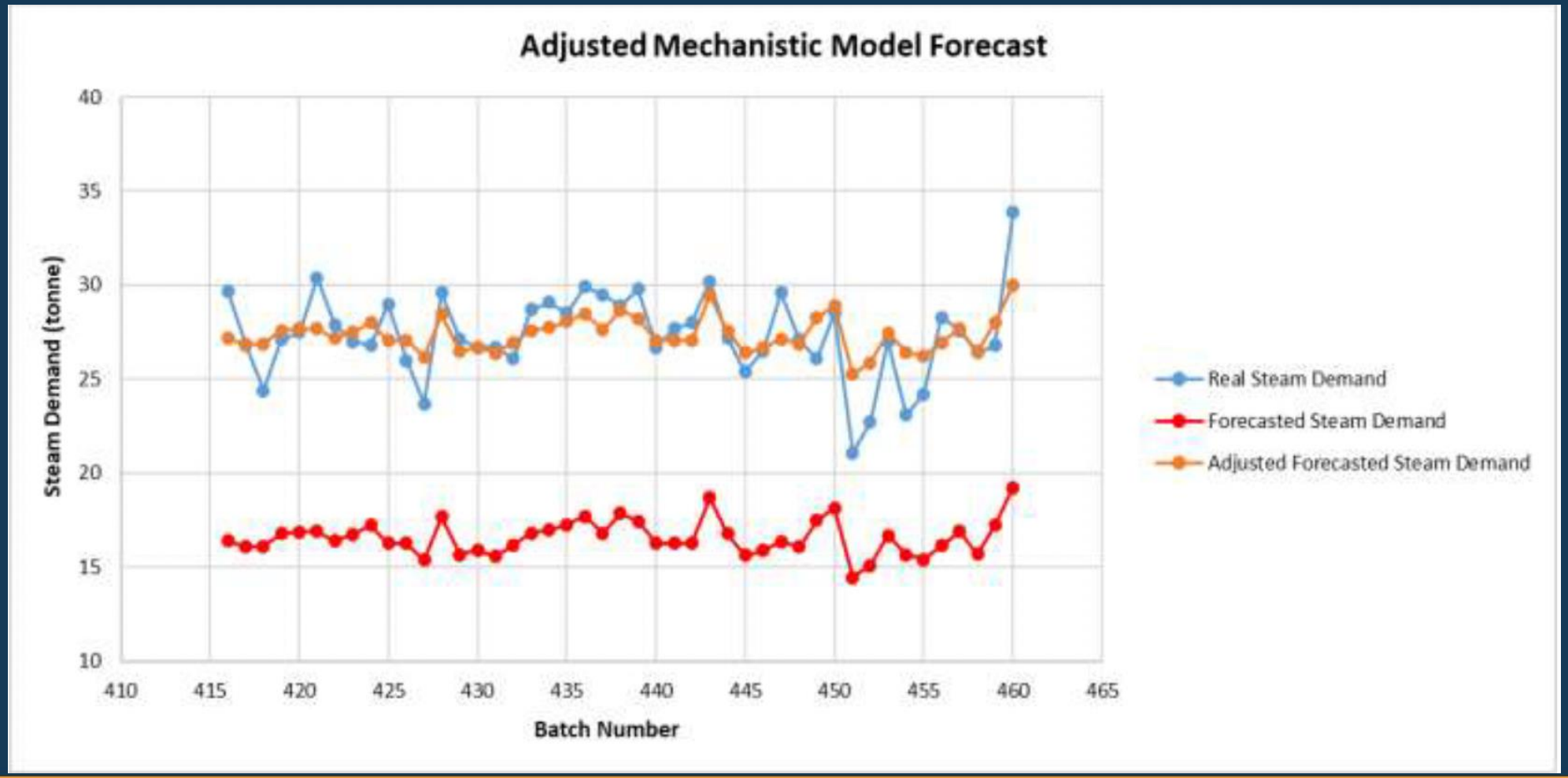
$$\frac{d\theta}{d\tau} = \frac{\sigma + \delta\rho}{1 + 0.47\alpha}$$

$$\theta = \sigma\tau - Da\delta(e^{-Da\tau} - 1)$$

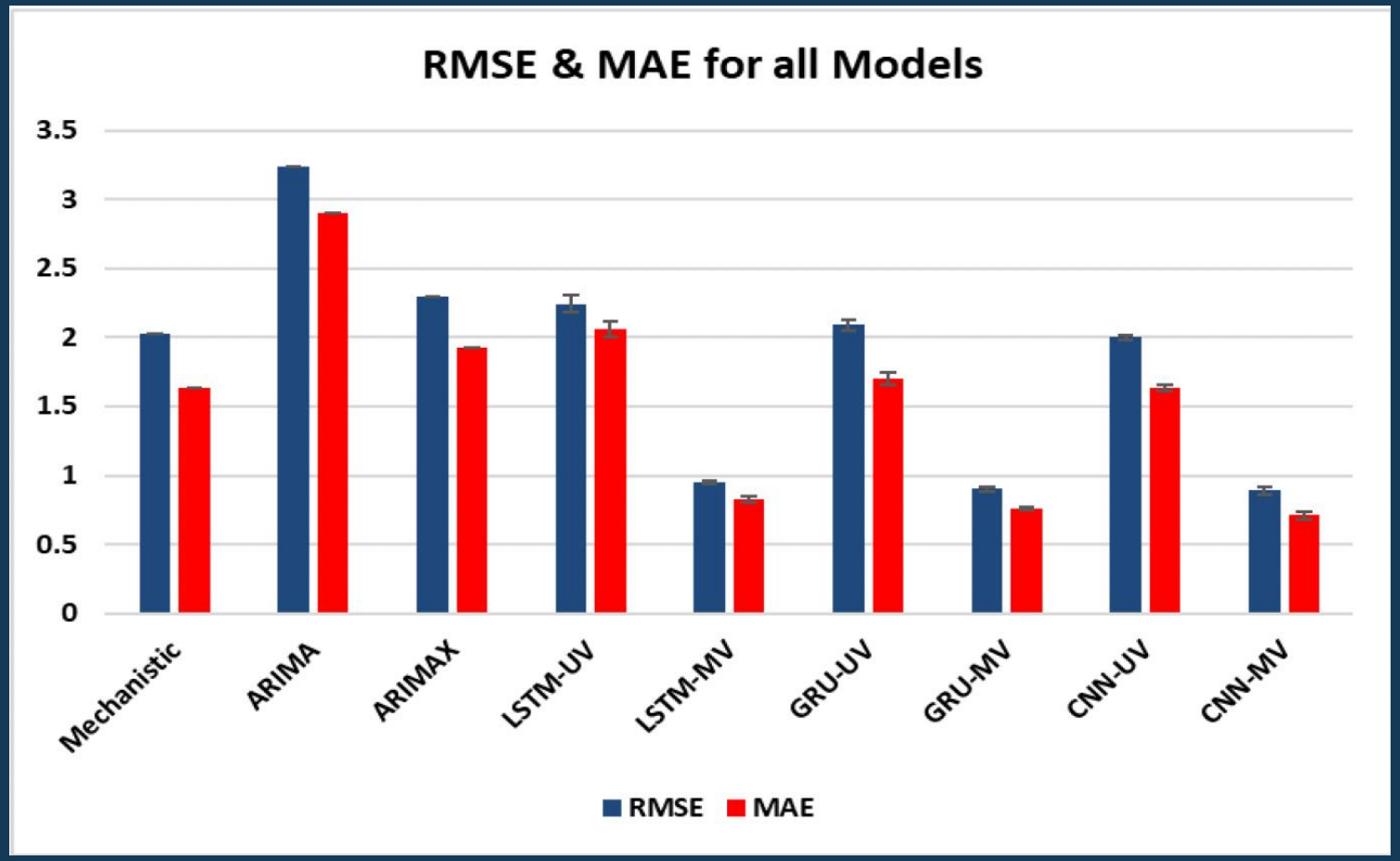
Steam demand forecasting



Steam demand forecasting



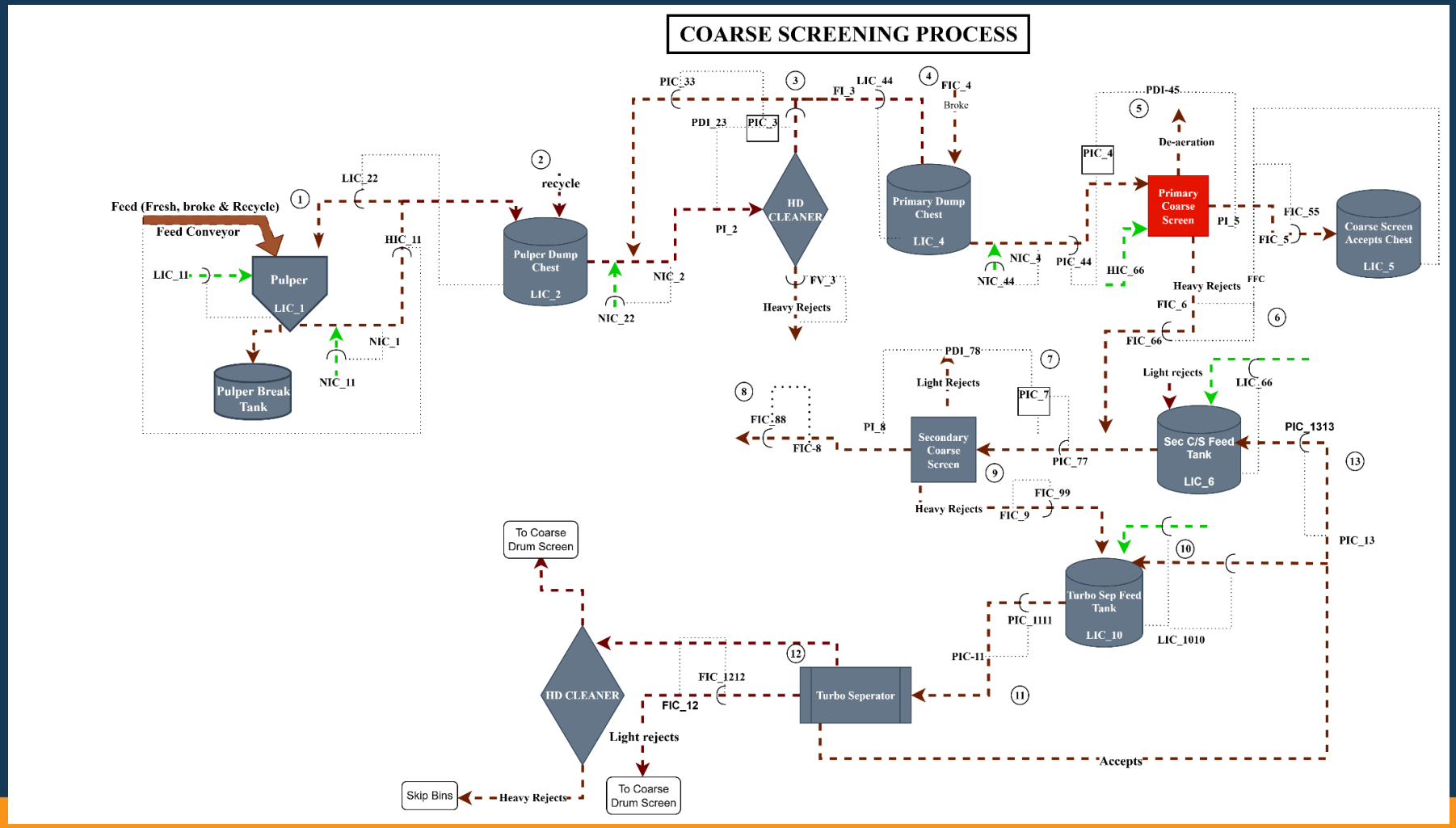
Steam demand forecasting



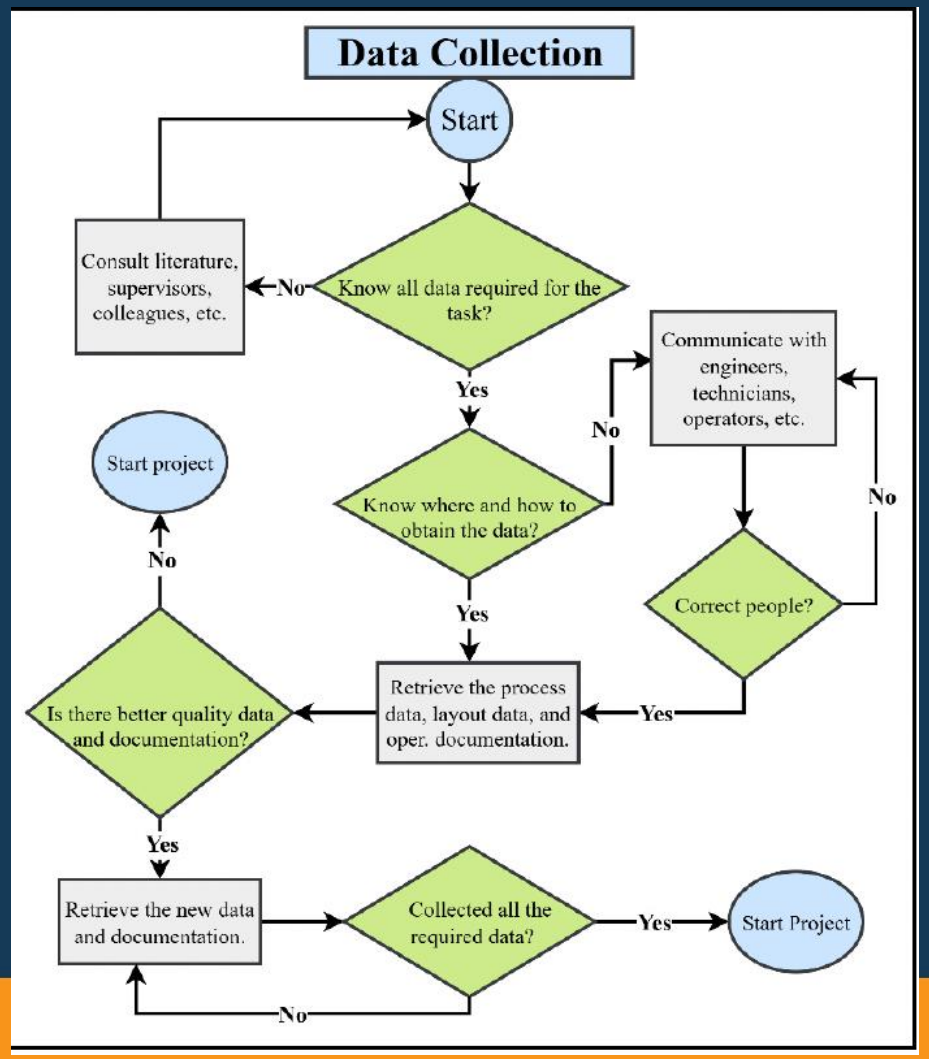
Steam demand forecasting

MODEL	RMSE	MAE	DEGREES OF FREEDOM
Mechanistic	2.03	1.63	3
ARIMA	3.24	2.90	5
ARIMAX	2.29	1.93	8
LSTM-UV	2.24 ± 0.062	2.06 ± 0.056	17425
LSTM-MV	0.95 ± 0.013	0.83 ± 0.022	17681
GRU-UV	2.09 ± 0.041	1.70 ± 0.041	13393
GRU-MV	0.90 ± 0.012	0.76 ± 0.014	13585
CNN-UV	2.00 ± 0.020	1.63 ± 0.023	1681
CNN-MV	0.89 ± 0.027	0.71 ± 0.024	1745

Fault detection in recycled fibre processing



Fault detection



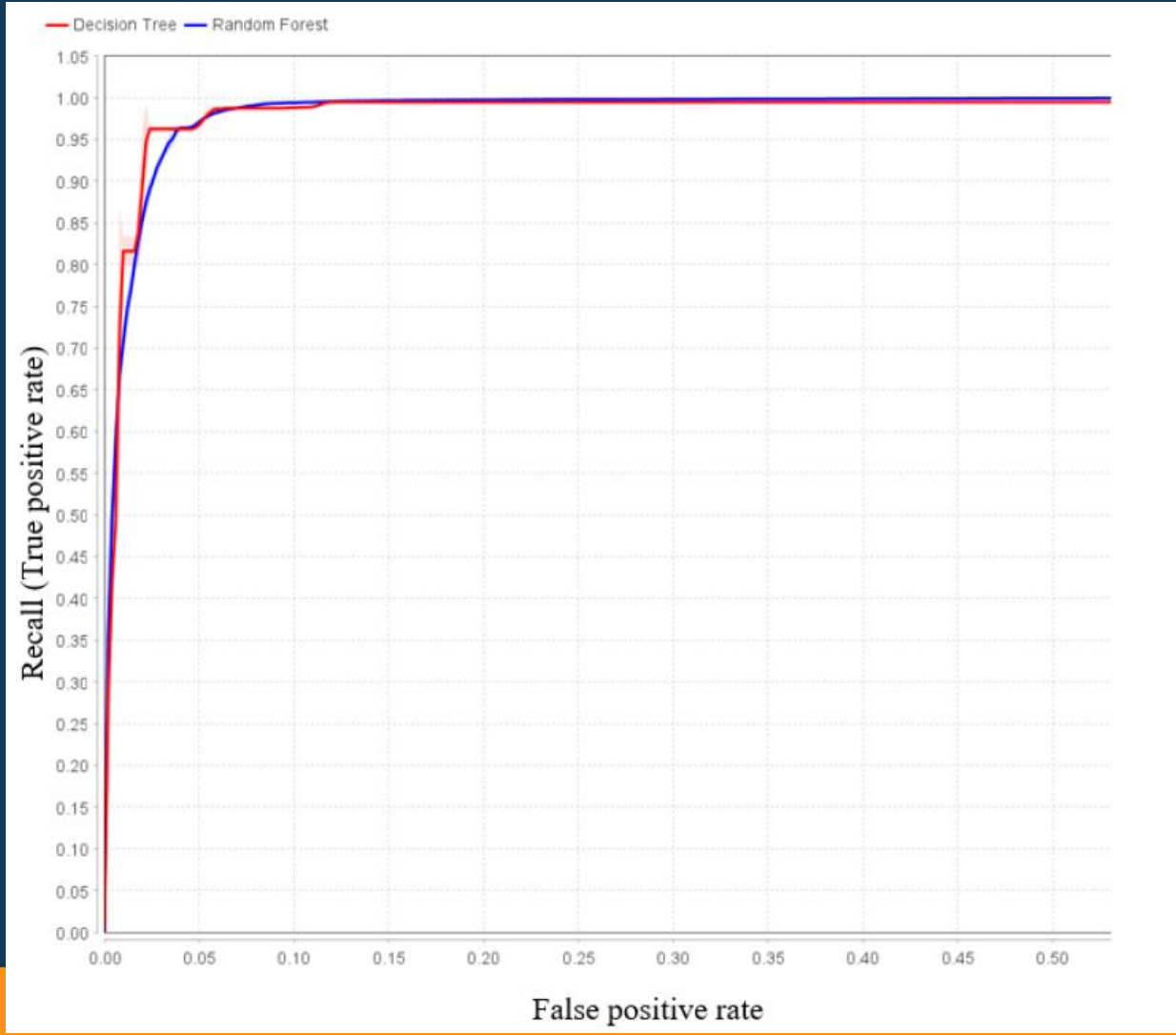
Fault detection-course screening

Variable	Normal Operation	Clogging Indication
PDI_45	< 40 Kpa	P > 40 Kpa
PI_5	180 Kpa average	P < 90 Kpa
HIC_66	30% average	Opens up to 70%
FIC_66	32% average	Opens up to 100%

Fault detection: course screening

	Model	Performance Metric	Value	Optimisation ease
1	Decision Tree	FCP	99.54%	Simple
		Recall	99.41%	
2	RF	FCP	99.44%	Simple
		Recall	99.35%	
3	GBDT	FCP	99.54%	Moderate
		Recall	99.77%	
4	SVM	FCP	99.50%	Complex
		Recall	99.17%	
5	NB	FCP	89.79%	Simple
		Recall	98.50%	
6	FFBPNN	FCP	98.80%	Complex
		Recall	99.67%	

Fault detection



Work moving forward

- The work performed by the DANCE group aims to support the South African process industry in its efforts to improve and optimise its processes.
- A variety of techniques are used to solve real-world problems. These include machine learning, first principles modelling, and novel control techniques.
- The work is enabled by the provision of datasets from industry, a key factor in enabling research in areas of specific interest.
- To continue on our upward path, we welcome discussion with the industry on the provision of datasets that will assist in the research of applied topics.

Key conclusions from DANCE research

- Tools are mature and easily available to allow the use of historical data for modelling, optimisation and control
- However the ease of use of statistical and modelling tools can lead to models that have excellent metrics but little function
- Data cannot replace engineering knowledge when developing models
- Availability of data but lack of capacity in industry provides an excellent opportunity for academic-industrial partnerships.

Acknowledgements

- Our industrial partners, Valterra, Implats, Harmony, Mt Isa, Mpact, Sappi,
- Students: Ms Trishkaya Harisunker, Ms Katlego Ramodike, Darren Bouttell, Tebogo Motsa, Davies Agommuoh, Madimetja Mashakoe.

THANK YOU!


DANCE
Data Analytics, Numerics and
Control Engineering Group


CHEMET
SCHOOL OF CHEMICAL AND METALLURGICAL ENGINEERING