# Machine Learning for Condition-Based Flowmeter Management


Kojo Sarfo Gyamfi

PhD Student

Supervisors:

Dr. James Brusey, Prof. Andrew Hunt, Prof. Elena Gaura


19 JULY 2016


Faculty of Engineering and Computing

Coventry University

# Content

# Introduction

- Flowmeters
- Subject to deviation: Waxing, Noise, Meter misalignment, etc.
- Problem of incorrect measurement – high flowrate attracts high tax liabilities.

- Recalibration: time of operation.
- Two main problems:

1) Malfunctioning meter before schedule.
2) Perfectly operating meter at schedule.

e.g. £30,000 for USM recalibration [5]

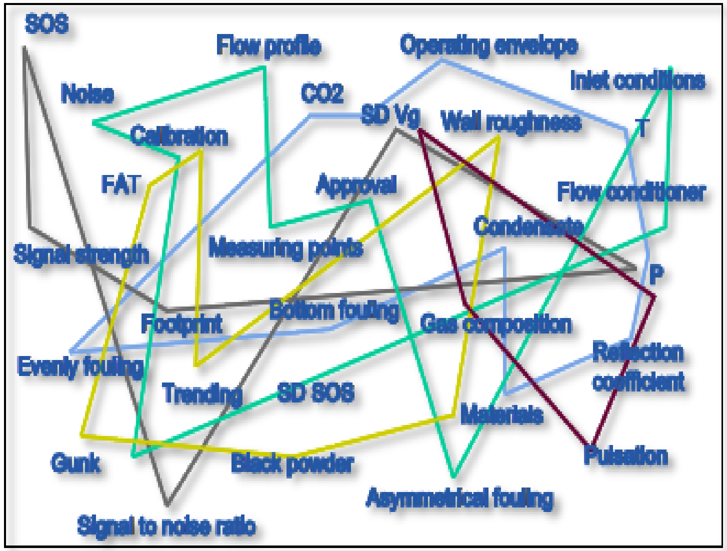
USM Flowmeter

- Condition-based flowmeter management (CBM).


- [5] Hall ,J., Zanker, K. And Kelner, E.

# Introduction

- CBM: condition of meter has to be known at all time
- Diagnostic parameters: USM, Coriolis: varied, sophisticated relationship




[2]

- Machine Learning:
- A pattern between diagnostic parameters and meter health state
- The exact physical or mathematical relationship is not known.
- There is data available.


- [2] Marcel J.M. Vermeulen, Jan G. Drenthen, Hilko den Hollander

# The Learning Problem

- What do we want to learn?

1) **What is wrong with the flowmeter?** Classification
2) **Does it affect the measurement integrity?** Regression
3) When will a recalibration of the meter be necessary?

Case Study: 4-path USM

**Classification**

4 Classes

1) Baseline, **B**
2) Gas Injection, **G**
3) Misalignment, **M**
4) Waxing, **W**

**Regression**

Estimation of measurement error $e$

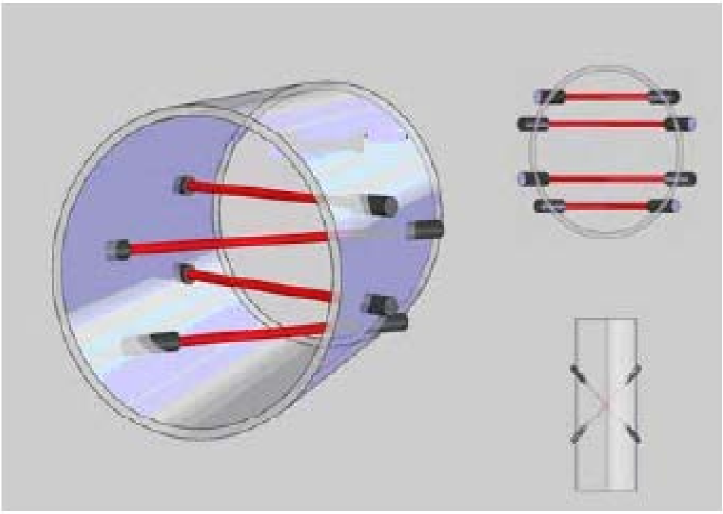
# The Learning Problem

- Dataset obtained from NEL: Full bore, Reduced bore USM

  - Profile factor
  - Crossflow
  - Speed of Sound
  - Performance
  - Transit time
  - Symmetry
  - Flow velocity
  - Signal Strength
  - Gain

$$\boldsymbol{d} = [p_1, p_2, \dots, p_M]$$

$$\mathcal{D} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \dots, \boldsymbol{d}_N]$$



- 361 data points of 48 dimensions each
- Monitoring of individual parameters, checking threshold violations [2]


- [2] Marcel J.M. Vermeulen, Jan G. Drenthen, Hilko den Hollander

# The Learning Problem

- Training and Testing

$$\mathcal{D} \rightarrow [\mathcal{D}_{train}, \mathcal{D}_{test}]: \quad [75\%, 25\%]$$

Training: $\mathcal{D}_{train}$

| Parameters | Label |
|---|---|
| $\boldsymbol{d}_1 = [p_1, p_2, \dots, p_M]$ | W |
| $\boldsymbol{d}_2 = [p_1, p_2, \dots, p_M]$ | B |
| $\boldsymbol{d}_3 = [p_1, p_2, \dots, p_M]$ | G |
| . | |
| . | |
| . | |
| $\boldsymbol{d}_{N_{train}} = [p_1, p_2, \dots, p_M]$ | M |

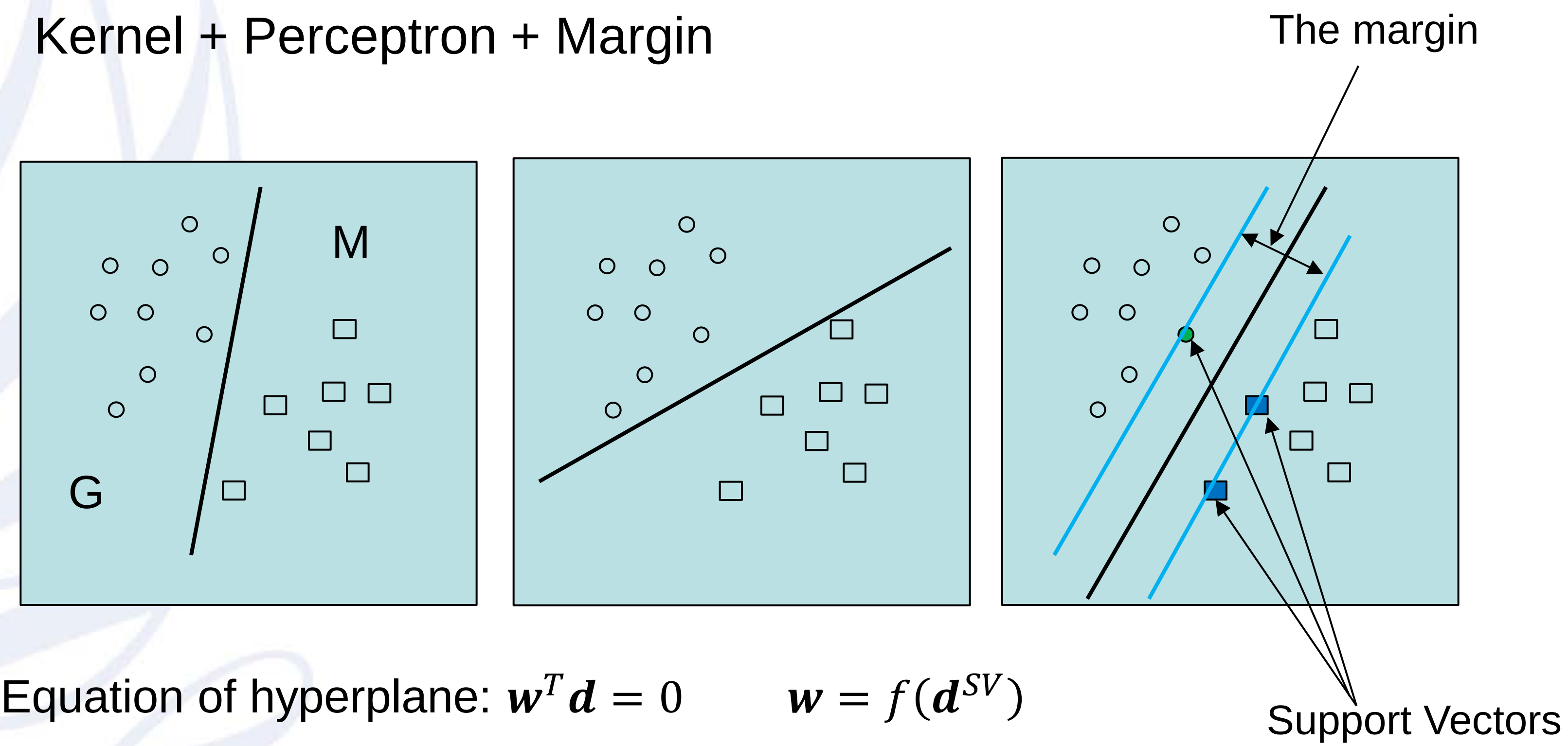
Hypothesis : $g(\boldsymbol{d})$

Target function : $h(\boldsymbol{d})$

Testing: $\mathcal{D}_{test}$

# Support Vector Machines

- Kernel + Perceptron + Margin




Equation of hyperplane: $\boldsymbol{w}^T \boldsymbol{d} = 0$ $\boldsymbol{w} = f(\boldsymbol{d}^{SV})$

SVM solves for $\boldsymbol{w}$ that maximizes the margin

The fewer the support vectors, the better the generalisation [9].

- [9] Caltech (2016) ,Learning From Data, [online] available from <http://work.caltech.edu/lectures.html> [3 April 2016]

Coventry University

# Support Vector Machines

- Multiclass Classification

One vs All

**B** vs **G, M, W**
**G** vs **B, M, W**
**M** vs **B, G, W**
**W** vs **B, G, M**

$K$ classifiers: 4

Highest score wins

Classifiers may yield different confidence values [8]

One vs One

**B** vs **G**
**B** vs **M**
**B** vs **W**
**G** vs **M**
**G** vs **W**
**W** vs **M**

$K(K-1)/2$ classifiers: 6

Majority wins

No clear majority! [8]


- [8] Bishop, Christopher M.

# Error Estimation

- Measurement error correlelates with diagnostic parameters.

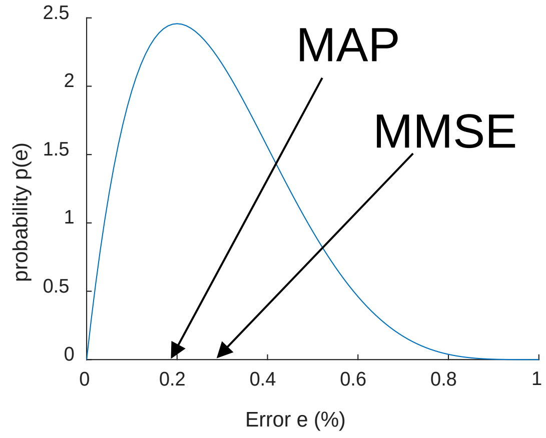
Joint distribution: $p(e, \boldsymbol{d})$

- Knowing $\boldsymbol{d}$ what is $e$ likely to be?

- Two estimates:

MAP estimate: $\arg\max p(e|d)$

MMSE estimate: $\arg\max E(\boldsymbol{e}|d)$



# Results: Classification

- **Reduced Bore**: Confusion Matrices

Predicted Labels

True Labels

| | B | G | M | W |
|---|---|---|---|---|
| **B** | 12 | 0 | 6 | 0 |
| **G** | 0 | 5 | 1 | 1 |
| **M** | 0 | 1 | 14 | 0 |
| **W** | 1 | 0 | 0 | 8 |

90.7 % accuracy

One vs All SVM

| | B | G | M | W |
|---|---|---|---|---|
| **B** | 12 | 0 | 0 | 0 |
| **G** | 0 | 6 | 0 | 1 |
| **M** | 0 | 0 | 15 | 0 |
| **W** | 0 | 0 | 0 | 9 |

97.7 % accuracy

One vs One SVM

# Results: Classification

- **Full Bore**: Confusion Matrices

Predicted Labels

True Labels

| | B | G | M | W |
|---|---|---|---|---|
| **B** | 11 | 0 | 2 | 0 |
| **G** | 0 | 5 | 0 | 0 |
| **M** | 1 | 0 | 13 | 0 |
| **W** | 0 | 0 | 0 | 13 |

93.3 % accuracy

One vs All SVM

| | B | G | M | W |
|---|---|---|---|---|
| **B** | 10 | 0 | 3 | 0 |
| **G** | 0 | 4 | 1 | 0 |
| **M** | 1 | 0 | 13 | 0 |
| **W** | 0 | 0 | 0 | 13 |

88.9 % accuracy

One vs One SVM

# Results: Classification

- **Full Bore + Reduced Bore**: Confusion Matrices

Predicted Labels

True Labels

| | B | G | M | W |
|---|---|---|---|---|
| **B** | 23 | 1 | 2 | 0 |
| **G** | 0 | 14 | 0 | 0 |
| **M** | 2 | 0 | 23 | 0 |
| **W** | 0 | 0 | 0 | 25 |

94.4 % accuracy

One vs All SVM

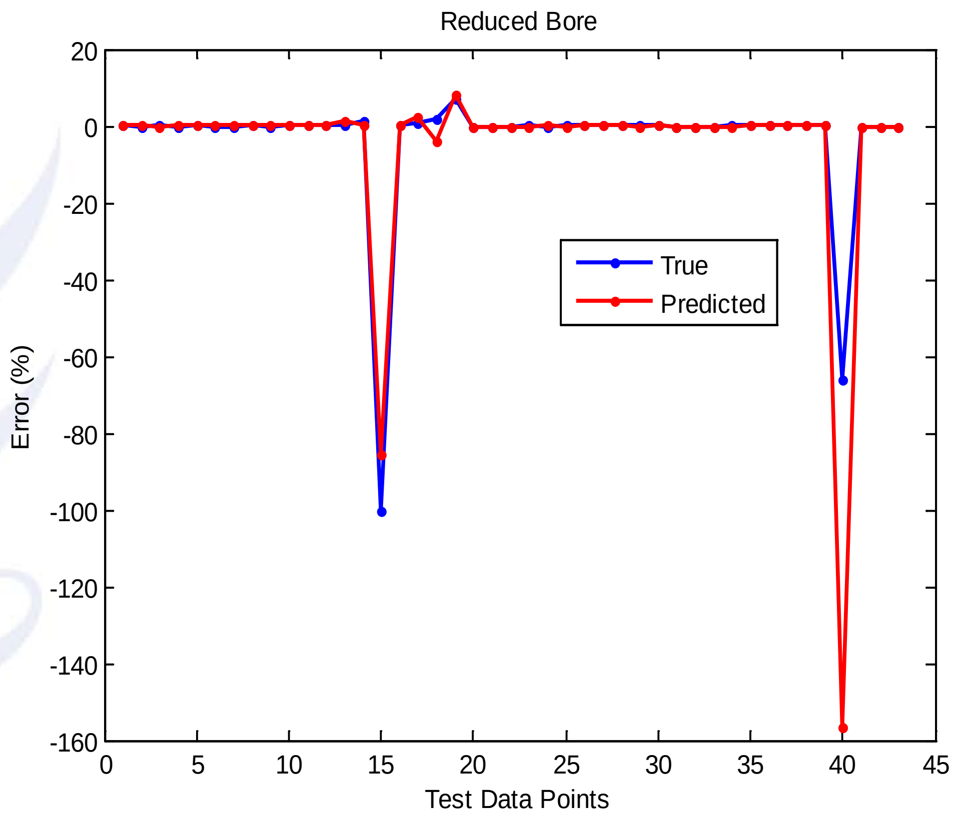
| | B | G | M | W |
|---|---|---|---|---|
| **B** | 22 | 1 | 3 | 0 |
| **G** | 0 | 13 | 1 | 0 |
| **M** | 3 | 1 | 21 | 0 |
| **W** | 1 | 0 | 0 | 24 |

88.9 % accuracy

One vs One SVM

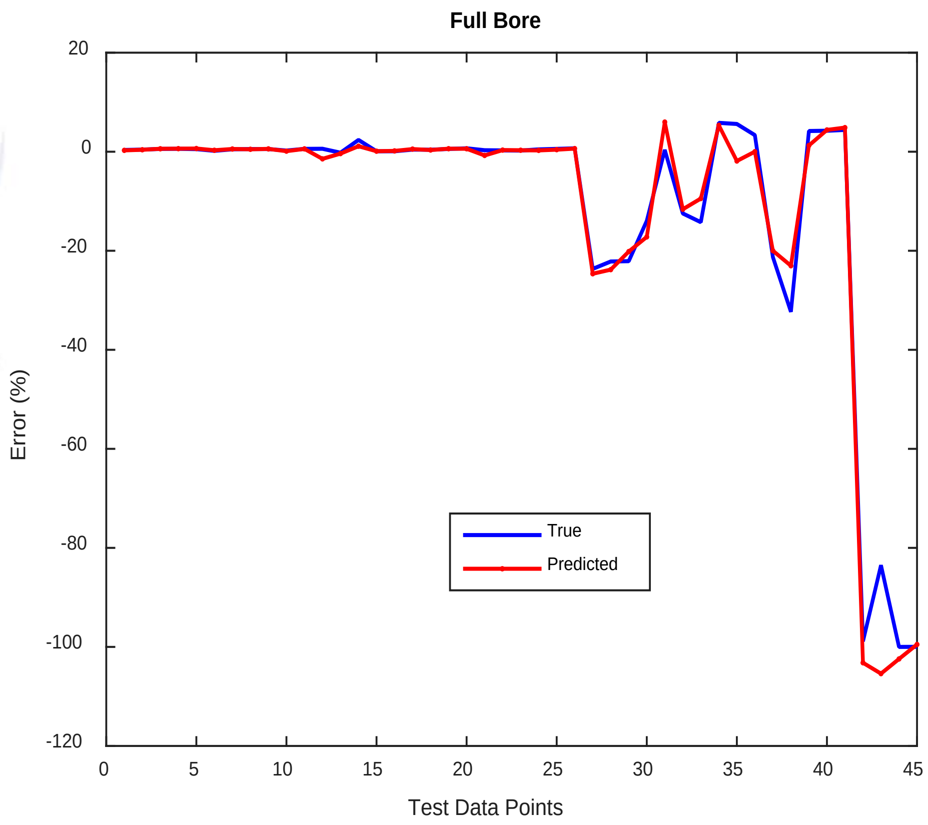
# Results: Error Estimation

- **Reduced Bore**



# Results: Error Estimation

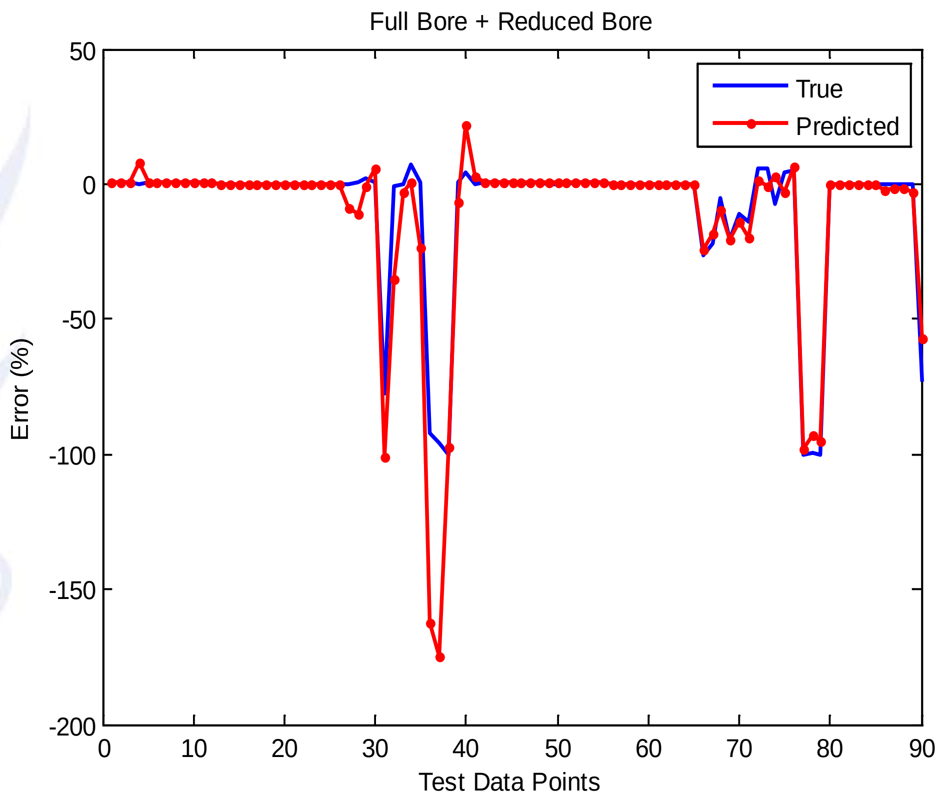
- **Full Bore**




Coventry University

# Results: Error Estimation

- **Full Bore+Reduced Bore**



# Conclusion

- There is a relationship between meter state, measurement error and diagnostic parameters.

- ML techniques may be used to learn this relationship, e.g. SVM, ANN, RBF, Bayesian Methods.

- ML techniques eliminate the need for end-user expertise, which may not always be available.

- Moreover, prediction of calibration frequency is possible with ML
  - Demands a lot of time-series diagnostic data.

# References


- [1] National Measurement System, "Good Practice Guide: The Calibration of Flow Meters”
- [2] Marcel J.M. Vermeulen, Jan G. Drenthen, Hilko den Hollander, “Understanding Diagnostics and Expert Systems in Ultrasonic Flowmeters”
- [3] National Measurement System, “Application of Ultrasonic Flowmeter Diagnostics”
- [4] B.Schölkopf et al., “Comparing support vector machines with Gaussian kernels to radial basis function classifiers,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2758-2765, Nov. 1997.
- [5] Hall ,J., Zanker, K. And Kelner, E., “When Should a Gas Ultrasonic Meter be Recalibrated,” 28th International North Sea Flow Measurement Workshop, 2010
- [6] Chapelle, Olivier; Schölkopf, Bernhard; Zien, Alexander, “*Semi-supervised learning”* MIT Press. (2006).
- [7] Hsu, C.-W., Lin, C.-J., “A Comparison of Methods for Multi-Class Support Vector Machines,” IEEE Transactions on Neural Networks, vol. 13, no. 2, pp. 415-425, Mar. 2002.
- [8] Bishop, Christopher M. (2006). Pattern Recognition and Machine Learning. Springer.
- [9] Caltech (2016) ,Learning From Data, [online] available from <http://work.caltech.edu/lectures.html> [3 April 2016]


Coventry University