Monitoring the Cleaning of Food Fouling in Pipes using Ultrasonic Measurements and Machine Learning

J. Escrig[1,2], E. Woolley[3], A. Simeone[4] & N. J. Watson[1*]

[1] Food, Water, Waste Research Group, Faculty of Engineering, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

[2] i2CAT Foundation, Calle Gran Capita, 2 -4 Edifici Nexus (Campus Nord Upc), 08034, Barcelona, Spain

[3] Wollfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough, LE11 3TU, UK

[4] Intelligent Manufacturing Key Laboratory of Ministry of Education, Shantou University, Shantou 515063, China

[*] Corresponding Author.
Address: Dr Nicholas James Watson. Room B16, Coates Building, University Park, University of Nottingham, Nottingham, Nottinghamshire, NG7 2RD, UK.
Email: nicholas.watson@nottingham.ac.uk
Phone: (+44) 115 74 84848

Abstract:

Food and drink production equipment is routinely cleaned to ensure it remains hygienic and operating under optimal conditions. A limitation of existing cleaning systems is that they do not know when the fouling material has been removed so nearly always over-clean, incurring significant economic and environmental costs. This work has studied the use of ultrasonic measurements and a range of different machine learning classification methods to monitor the fouling removal of food materials in plastic and metal cylindrical pipes. The experimental results showed that the developed techniques could predict the presence of fouling with prediction confidence as high as 100% for both plastic and metal pipes. The sensor technique performed marginally better in the plastic pipe and similar performance was found for the all of the machine learning methods studied. This work has demonstrated the potential of low-cost ultrasonic sensors to monitor and therefore optimise cleaning processes within pipes. It is discussed how new data set labelling strategies will be required for the techniques to be used effectively within production environments.

1.   Introduction

Cleaning of processing equipment is essential for manufacturing sectors such as food and drink, pharmaceutical and Fast Moving Consumer Goods (FMCG). Cleaning is required to ensure equipment remains hygienic and operating under optimal conditions, in addition to reducing the possibility of cross contamination. This is especially important when the processing equipment may be used to manufacture different products, some of which may contain allergens. Cleaning the internal surfaces of processing equipment is generally an autonomous task performed by Clean-in-Place (CIP) systems. The benefits of CIP systems is that they do not require any equipment disassembly or operator input once they have begun. They also produce a standardised cleaning which is routinely validated to determine surface cleanliness using Adenosine Triphosphate (ATP) and growth and microbial enumeration tests (Fratamico, Annous, & Guenther, 2009). Clean-in-place systems use a combination of mechanical force (caused by the flowing liquid), chemicals, temperature and time to clean the equipment and generally feature the following stages: pre-rinse, chemical cleaning, post-rinse, sanitisation and final rinse (Fryer, Christian, & Liu, 2006).

Although CIP systems are industry standard they are not perfect and opportunities remain to improve their operation. Due to the large amounts of water, chemicals and energy CIP systems utilise they have a significant environmental impact. It has been reported that cleaning is responsible for 30% of the energy used in dairy processing (Eide, Homleid, & Mattsson, 2003), and approximately 35% of the water used in beer production (Pettigrew, Blomenhofer, Hubert, Groß, & Delgado, 2015). Cleaning also takes a significant amount of time which affects the productivity of a manufacturing facility as any time spent cleaning cannot be used to manufacture products. Most manufacturers will produce a range of different products on the same processing equipment, yet CIP systems are generally only developed for one or two different formulations which are known to be the most difficult to clean. This results in significant over cleaning for most other products adding to their environment and economic cost. Therefore, there is a need for new technologies to optimise current industrial cleaning processes.

Manufacturing is experiencing the fourth industrial revolution (Industry 4.0) which is underpinned by the enhanced collection and use of data within production environments, in addition to other industrial digital technologies such as robotics and the industrial internet of things. Industry 4.0 is been increasingly adopted in the discrete manufacturing sector (e.g. automotive) but less so in process manufacturing sector (e.g. food and drink) due to the lack of cost effective sensing technologies

capable of monitoring processes within production environments (Alexander Lewis Bowler, Bakalis, & Watson, 2019; Fisher et al., 2018). Research has been performed investigating the use of sensor technologies to monitor cleaning processes. Some of this work has focussed on monitoring the cleaning fluids as they exit the equipment using different sensing technologies (Lyndgaard, Rasmussen, Engelsen, Thaysen, & Van Den Berg, 2014) (Berg, Ottosen, Berg, & Ipsen, 2017) (Van Asselt, Van Houwelingen, & Te Giffel, 2002). Although these techniques are useful and have led to commercial technologies they are only capable of detecting fouling which has transferred to the cleaning fluid and not that which remains on the internal equipment surfaces. This may lead to inaccurate indications of the level of equipment cleanliness. Different sensor techniques have been studied to directly detect the presence of fouling on surfaces including electrical (e.g. (E. Wallhäußer, Hussein, & Becker, 2012) (X. D. Chen, Li, Lin, & Necati, 2003) (Guérin, Ronse, Bouvier, Debreyne, & Delaplace, 2007) (Tlili, Rousseau, Ben Amor, & Gabrielli, 2008)), acoustic (e.g. (Pereira, Mendes, & Melo, 2009) (P. M. Withers, 1996) (Úbeda, Hussein, Hussein, Hinrichs, & Becker, 2016)) and optical methods (e.g. (P. M. Withers, 1996) (Simeone, Deng, Watson, & Woolley, 2018) (Cheong, Kim, & Kim, 2017)(Simeone, Watson, Sterritt, & Woolley, 2016) (Tamachkiarow & Flemming, 2003)). Electrical methods generally measure changes in either electrical conductance or resistance between a pair of probes. Although sensitive to fouling material and changes in heat transfer within equipment they can be difficult to utilise in environments, such as cleaning, where turbulent flow is present. Acoustic methods are sensitive to vibrational changes within equipment. Although these sensors are non-invasive and sensitive to surface fouling (as fouling modifies vibration in a surface) they are difficult to utilise in industrial environments where many other vibrations exist. Optical methods can provide quantitative spatial and temporal information on the surface fouling as it is cleaned, so are an attractive technology. However, for optical techniques to operate effectively they must be able to view the surface been cleaned. With suitable lighting this is possible in large tanks but is extremely difficult in pipes of small diameter (e.g. < 25 mm) which are common in many industries.

Of interest in this research are Ultrasonic (US) techniques which use mechanical waves operating at frequencies above 20 kHz. Ultrasonic techniques have been previously used to detect fouling on surfaces and monitor its removal during cleaning processes (B. Chen et al., 2019; Escrig, Escrig, Woolley, Rangappa, Simeone, & Watson, 2019; Lohr & Rose, 2003; Úbeda et al., 2016; E. Wallhäußer et al., 2012; E. Wallhäußer, Hussein, Hussein, Hinrichs, & Becker, 2013; Eva Wallhäußer, Hussein, Hussein, Hinrichs, & Becker, 2011; Eva Wallhäußer et al., 2014; P. Withers, 1994; P. M. Withers, 1996). The majority of this previous work has focused on dairy fouling due to its industrial importance.

Ultrasonic methods and are an attractive sensing technology to detect fouling as the US waves are sensitive to changes in material properties of the systems in which they propagate. In addition

ultrasonic sensors are low cost and have the ability to operate non-invasively on opaque systems (Watson, 2015). However there are several limitations of ultrasonic methods including their sensitivity to temperature making measurements in industrial environments challenging. In addition they are a local measurement requiring careful consideration of where sensors should be placed or the use of multiple sensors.

With any sensing technique there is a need for data analysis methods which utilise the recorded sensor measurements and provide information on the system under investigation. Machine Learning (ML) is a popular method for sensor data analysis as the techniques do not require the development of physical models to account for complexities such as environmental effects (e.g. temperature and humidity). In addition ML models are known to improve as more or better data becomes available (Goodfellow, Bengio, & Courville, 2016). Classification ML methods have been used in combination with ultrasonic measurements to detect fouling on material surfaces (Wallhaußer, Hussein, Hussein, Hinrichs, & Becker, 2011), (Úbeda et al., 2016; E. Wallhäußer et al., 2013; Eva Wallhäußer et al., 2014). These works have focussed exclusively on dairy fouling and found that Artificial Neural Networks (ANN) and Support Vector Machines (SVM) can predict the presence of fouling in stationary and flowing systems with prediction accuracies as high as 99%. Although previous work has demonstrated the potential of the technique it has only been performed on flat test sections and not circular pipes, which are much more common in industrial environments. Ultrasonic techniques can be used successfully on cylindrical pipes for applications such as flow monitoring (Al-Aufi et al., 2019). However it is known that the amplitude of reflected US wave reduces as the curvature of a pipe increases (W. Wang, Liu, Li, Xu, & Zhang, 2019). This may limit the capability of US techniques to accurately detect the presence of fouling in industrial environments, especially when only a small degree of fouling is present.

This current work investigates the use of US sensors and ML classification methods to detect the presence of surface fouling in pipes with circular cross section geometry. Pipe sections of plastic and metal are studied for different industrially relevant food fouling materials and temperatures. A range of different ML classification methods are utilised to determine their effectiveness in predicating the presence of fouling from the US measurements.

2. Materials and Methods

2.1 Experimental measurements

Two circular pipes were used for the experiments in this work. The details of the pipes are in Table 1 and Figure 1. A 1" (25.4 mm) pipe connecter was bonded to one end of each pipe so that it could be

attached to a water supply. The purpose of studying pipes of different materials was to understand the effect that this has on the performance of the developed sensor technique.

Table 1: Properties of pipe sections used during experiments

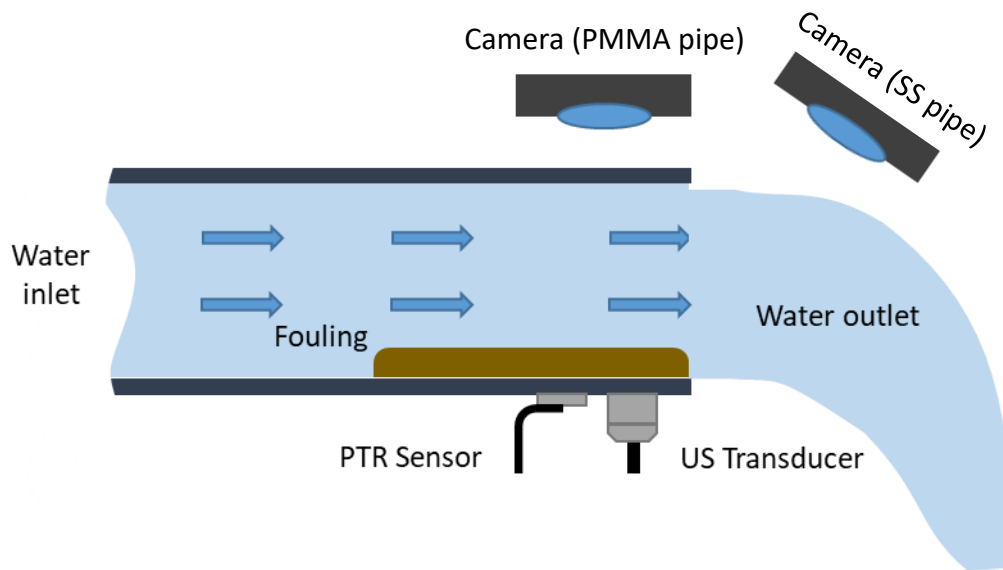|	Material of construction	Length (mm)	Internal Diameter (mm)	Wall thickness (mm)
Plastic pipe (PMMA)	Polymethyl methacrylate	340	20	2.2
Stainless steel pipe (SS)	Stainless steel	270	22	1.65

Figure 1: Experimental setup including sensor and fouling locations

2.1.1 Sensors

Ultrasonic transducers were attached to both pipe sections. A single transducer was used for each pipe section and operated in reflection mode. The transducers used were 2 MHz contact transducers manufactured by Yushi (model: 2P10N). The transducers were attached to the pipes using a permanent adhesive and were located approximately 30 mm from the opposite end of the pipe where the water supply was connected. During all experiments the pipes were orientated so that the transducer was located at the bottom (Figure 1). The US transducers were connected to a US box (Lecoeur Electronique®), which was utilised to excite the transducers and digitise the received US waves. The US box features an ultralow noise preamplifier, has a – 6 dB bandwidth of 500 KHz to 18 MHz, a 12 bit digitiser and provide up to 80 dB of gain.

To train the ML models with data from the US measurements a second method was required to determine if fouling was present in the pipe. In this work a camera was used for this purpose. It is worth noting that it would not be possible to use a camera in industrial environments as they would

need to be placed within the pipes and require suitable lighting. The camera utilised in this work was a Logitech® C270 3MP web camera and was located facing the open end of the pipe. It was positioned to provide the clearest view of the fouling material (Figure 1). For the SS pipe it was difficult to record clear images of the fouling due to light reflections within the pipe. Temperature was recorded using a RTD PT100 which was attached to a Pico Technology® PT-104 data logger. The US box, PT-104 data logger and web camera were all connected to a laptop. Instrumentation control and data collection were all performed via bespoke software developed in MATLAB®. During the experiments, US waveforms and temperature data was recorded every four seconds and images from the camera were recorded every 20 seconds. The US transdcuer was excited by a 200 volt square top pulse generated by the US box. The received signals were amplified by 18 dB and digitised at a sampling rate of 160 MHz. Once an US wave is generated in the transducer it will propagate through the pipe wall and become incident with the pipe/fouling layer interface. At this point a proportion of the wave will be reflected and travel back through the pipe wall before being detected by the same transducers. The propotion of the wave that is reflected is dependent on the acoustic impedance mismatch between the pipe wall material and the fouling material. One challenge when using US methods in thin plates, such as the experiment rig utilised in this work, is that the reflected signal may return to the transducer before the transmitted signal has fully left. This results in overlapping US waves. Although the majority of previous work utilising US to monitor fouling used test sections with sufficient thickness to separate reflected US signals in time (e.g. (Eva Wallhäußer et al., 2011)) it has also been shown that it is still possible to analyse US waves were overlapping has occurred (B. Chen et al., 2019). The ability for US techniques to operate on thin walled sections is essential for them to experience more wide-spread use in industrial environments.

2.1.3 Fouling material and experimental methods

Tomato paste and concentrated malt extract were used as the fouling media. Two different food materials were studied as it has been reported that different materials foul, and are cleaned from surfaces differently (Fryer & Asteriadou, 2009) (Cuckston, Alam, Goodwin, Ward, & Wilson, 2019). The tomato paste was Napolina® Double Concentrate Tomato Pure and the ingredients were: tomatoes, acidity regulator (citric acid). The concentrate malt was taken from a Coopers® Real Ale beer kit with the following ingredients: malted barley, hops, yeast and water. The fouling layer was created by depositing 15 grams of the material on the bottom of the pipe in the same axial and radial location of the US transducer (Figure 1). The fouling layer was then spread evenly with a spatula to form a uniform layer of approximately 5 mm thickness. It was left to dry for ten minutes before beginning the cleaning experiments.

Numerous cleaning experiments were performed to collect sufficient data to train and test the ML models. Cleaning was performed with water at two different temperatures (12℃ and 45℃) for both fouling materials. Full details of the cleaning experiments and how these were split into training and test data can be seen in Table 2.

Table 2: Cleaning experiments used for training and test data sets

PMMA Pipe						
Experiment data usage	**Material**	**Temperature (℃)**	**Flow rate (l/min)**	**Repetitions**	**Number of recorded US waves**	**Camera images (target data)**
PMMA Training	Tomato	12	6	6	200	40
	Tomato	45	6	6	200	40
	Malt	12	6	6	200	40
	Malt	45	6	6	200	40
PMMA Test 1	Tomato	45	6	6	200	40
PMMA Test 2	Tomato	12	6	1	200	40
PMMA Test 3	Malt	45	6	1	200	40
PMMA Test 4	Malt	12	6	1	200	40
Stainless Steel Pipe						
SS Training	Tomato	12	6	8	200	40
	Tomato	45	6	6	200	40
	Malt	12	6	6	200	40
	Malt	45	6	6	200	40
SS Test 1	Tomato	45	6	1	200	40
SS Test 2	Tomato	12	6	1	200	40
SS Test 3	Malt	45	6	1	200	40
SS Test 4	Malt	12	6	1	200	40

To begin the experiments data collection was initiated and then water was supplied to the pipe at the specified temperature and flow rate (calibrated via preliminary experiments). Each experiment continued for approximately 15 minutes which was sufficient to remove the fouling material and record enough data of a fouled and clean pipe.

2.2 Machine learning

2.2.1 Input and target data selection method

K-best predictors/features is a common methods used to determine the most suitable input data points to use for a ML model. Within this work all sampling points within a recorded waveform could be potential input data points for the ML models. A signal sampling frequency of 160 MHz was utilised for the US wave acquisition so a typical recorded waveform would have approximately 17,000 potential data points. For all recorded waveforms it was found that the first part of the signal was saturated so were not used as potential input data points. A signal segmentation procedure was applied to remove the saturated component resulting in approximately 11000 remaining data points. All of these points could be used as potential inputs into the ML models. However, some of these data points may not provide key information and would increase the time require to train, validate and test the models. The optimal number of input data points (K) was determined using a grid search during the validation of the models and the relevancy of each data point was determined using an F-test (Polat & Güneş, 2009). The results from these searchers can be seen in Table 3 and 4 in the results section.

The ML models were developed using all the data for both materials at the two temperatures. Separate models were developed for the plastic and metal pipes. It was decided to combine the data from all experiments for a single pipe to ensure a sufficient amount of data was available for effective model training. The dataset was split for training and validation using the K-Fold method. The target dataset was constructed by assigning each US waveform the label of either 'clean' or 'dirty'. The US transducer has an active element area of 1 cm^2 and is only affected by fouling on the opposite side of the pipe where the transducer is located. A digital image was acquired every 20 seconds during cleaning. However, US waveforms were recorded every four seconds so there were 5 recorded US waveforms for every image. To ensure the US waveforms were labelled as either clean or dirty, if any fouling was observed at the active element area of the image the five US waveforms recorded in the following 20 seconds were labelled as 'dirty'. If no fouling was detected in the images the waveforms were labelled as clean. This labelling process may result in errors in identifying the precise time the pipes became clean which may have an effect of the classification results. The training and test datasets for both pipes are shown in Table 2.

2.2.2 Classification methods

In this work the ML algorithms used were K-Nearest Neighbour (KNN), SVM, and ensemble methods such as Random Forest (RF) and AdaBoost. Artificial Neural Networks (ANN) are a popular ML method but were not studied as they have been shown to perform worse than SVM in previous work using US measurements and ML to detect fouling (Eva Wallhäußer et al., 2014). A KNN classifier is a non-parametric method where the classification of a new instance has the same value as the training data points that are closer to this instance in the feature space. If k is equal to 1 it has the same value as the closest neighbour and if k=n the instance is classified after the vote of the n nearest neighbours (Altman, 1992). A SVM is a method based on statistical learning theory aimed at determining the location of decision boundaries yielding the optimal separation of classes (Vapnik, 1999), i.e. 'clean' or 'dirty' in this case. For binary classification problems in which the classes are linearly separable, the SVM selects from amongst the infinite number of linear decision boundaries, the one that minimizes the generalization error. Accordingly, the selected decision boundary will be the one that leaves the greatest margin between the two classes, i.e. the sum of the distances to the hyperplane from the closest points of the two classes (Boser, Guyon, & Vapnik, 2004). The resulting data points closest to the hyperplane are used to measure the margin; therefore such data points are defined "support vectors" (Vapnik, 1999). In case the two classes are not linearly separable, the SVM tries to find the hyperplane that maximizes the margin while minimizing the misclassification errors. Support Vector Machines can also be extended to deal with nonlinear decision surfaces, by projecting the input data onto a high-dimensional feature space using kernel functions and formulating a linear classification problem in that feature space (Boser et al., 2004).

Decision trees are a hierarchical model that can be used for classification. The instances to be classified advance through the different nodes of the tree until they finish in a labelled leaf node. The path through the nodes that they follow depends on a series of conditions that the features of the instances need to meet. The training of a decision tree classifier consists of choosing the conditions that perform the best classification of the instances. Several algorithms exist to set up these conditions (Hastie, Tibshirani, & Friedman, 2009). The depth of the tree is the maximum number of splits that a path inside the tree can follow. A deep tree can classify very accurately all the training data, but it often presents problems of overfitting when tested with new data. Another problem is that decision trees are not very robust because a small change in the data can produce a completely different tree and very different predictions.

2.2.2.1 Ensemble Learning

Ensemble learning aims at developing a population of base learners (such as decision trees) from the training data, and then combines them to form a composite prediction model. It consists of building

multiple different decision tree models from a single training data set by recursively adopting multiple bootstrapped data subsets and averaging the models. Each tree is built independently to the others (Ruppert, 2004). Random Forest (RF) is a bagging-based method that produces multiple trees by selecting only some of the instances of the training data and considering only some of the features of these instances to build the decisions trees. For each new tree, the instances and the features of these instances used for the training of the tree are chosen randomly. The classification of a RF is the result of the vote of the solutions given by every single decision tree. Random forests have been shown to be very robust compared to simple decision trees (Alpaydin, 2014).

Boosting is another ensemble method that refers to the problem of building a strong learner out of a collection of weak learners; i.e. learners whose predictive accuracy is only slightly better than random guessing (Sabzevari, Martínez-Muñoz, & Suárez, 2018). Adaptive boosting (AdaBoost) is a second level ML method that combines the predictions of simple algorithms, in this case decision trees, to give a final prediction. This is calculated as the weighted sum of the predictions of the first level algorithms:

$$F(x) = \sum_{i=1}^{N} w_i \, f(x) \quad (2)$$

Where $f(x)$ is the function of a simple algorithm that gives a prediction of an x input, w_i is the weight of each algorithm considered in the ensemble, N is the total number of algorithms used, and *F(x)* is the final prediction. AdaBoost algorithms calculate the weights, w_i , iteratively to reduce the error of the prediction. The first level algorithms used in this work were decision trees of depth equal to 1. They were trained independently using random parts of the training data set.

2.2.3 Selection of parameters, grid search and k-fold validation

There are several parameters of the training algorithms that need to be selected for the different classification methods. These are: the number of input data points, the maximum depth of the trees in the RF algorithms, the number of neighbours to consider in the KNN algorithm and the learning rate applied in the iteration of the AdaBoost. These parameters were optimized through a grid search and a K-fold validation method. The number of input data points (taken from the US waveforms) introduced into the grid search were values from 25 to 1200 in steps of 25 for all of the algorithms. The maximum depths of the RF algorithm considered in the search were between 3 and 10 inclusive. The number of neighbours in the KNN algorithm used in the validation ranged from 1 to 200. Finally, the learning rates studied in the AdaBoost were 0.5, 0.75 and 1. The k-fold validation was performed by dividing the training data into 5 folds. The parameters of the grid search that scored higher in the k-fold validation were included in the different models. Other parameters were not included in the grid search because adding more dimensions to the grid increases the calculation time exponentially. The number of trees produced in the RF and AdaBoost algorithms was set to 500 and the minimum number of instances in each leaf during the training of the trees was fixed to 10. The classification

models developed in this work were tested on data not used for training or validation (see Table 2). To determine the accuracy of the different algorithms studied the results from the different classification algorithms were compared to the images recorded at the same time as the US measurements.

3   Results and Discussion

3.1 Input data selection

Figure 2 (a) shows a received US waveform recorded during the experiments. The non-saturated part of the waveform is coloured black and is located from approximately the 300th data point. The red dots show the 200 selected wave data points (From the non-saturated wave) used as inputs in the KNN model. These input data points were predominately located between the US wave sampling points 300 to 800 (Figure 2 (b)).
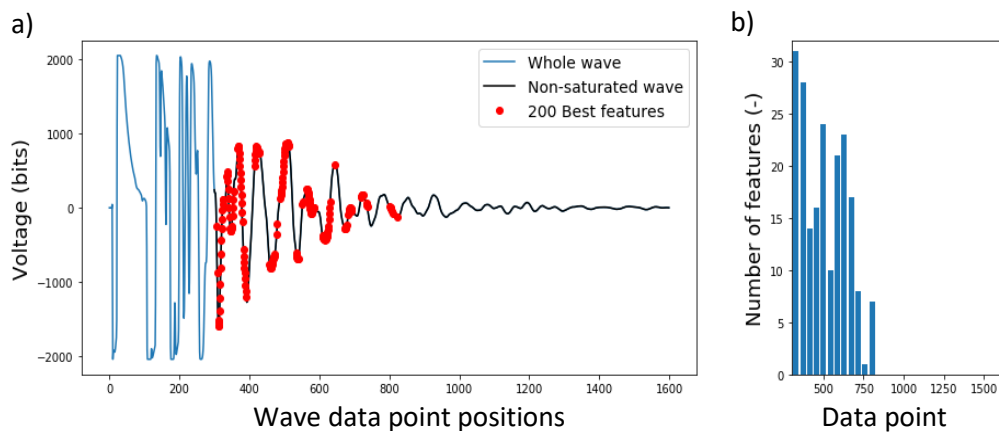

Figure 2: a) Recorded waveform from the PMMA pipe with location of best input data points. b) Cumulative location of the best input data features.

Figure 3 displays how the number of input data points extracted from the US waves affects the validation score for the different classification methods utilised for key model parameters. For all of the different classifiers studied, a lower number of input data points generally resulted in a higher Cross Validation (CV) score. This was especially the case for the SVM (Figure 3 b). For consistency the number of input data points which resulted in the highest CV score was used for all models. For the different model parameters studied in the grid search, in general the lower values resulted in a higher CV score. The optimal number of input data points and the values utilised for other model parameters can be seen in Table 3. The values with an asterisk denote ones where the values were selected by the researchers.
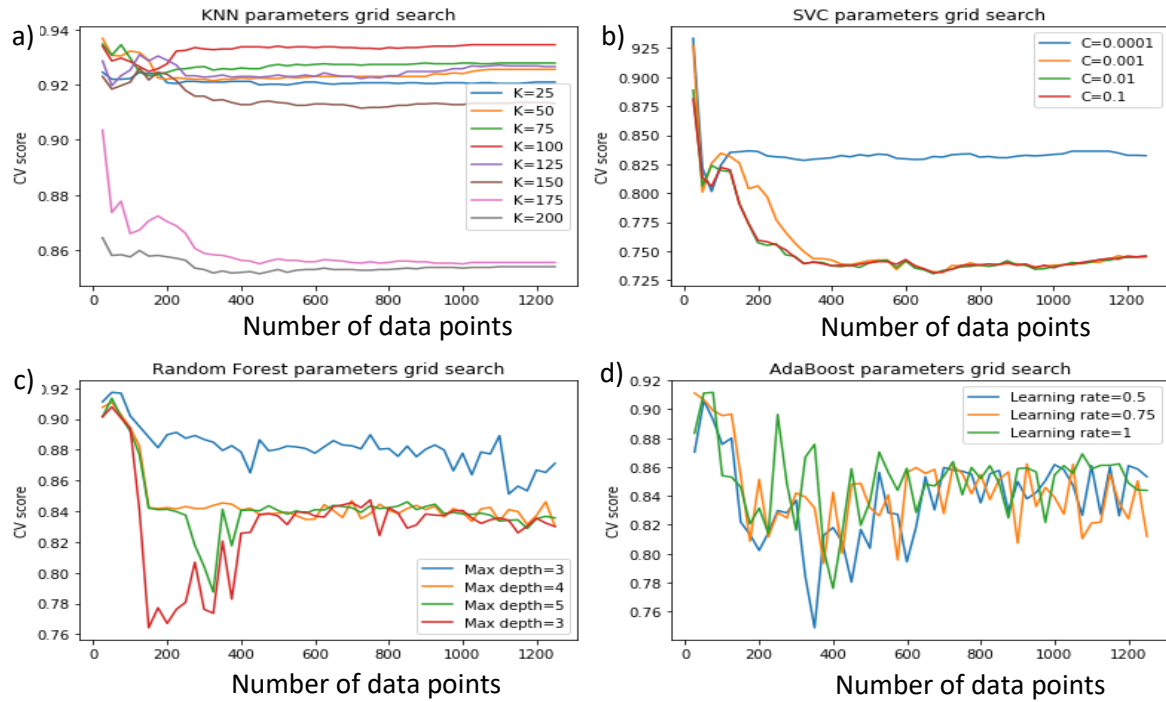
Figure 3: Cross validation (CV) scores for varying number of input data points and other key model parameters for the different modelling methods used in the PMMA pipe. a) KNN b) SVM c) RF d) AdaBoost.

Table 3: Classification model parameters for PMMA pipe

Model Parameter	KNN	SVM	RF	AdaBoost
Number of input data points	25	25	50	75
Number of neighbours	50	-	-	-
Number of trees	-	-	500*	500*
Maximum depth:	-	-	3	1*
Minimum number of instances in leafs	-	-	10*	10*
Learning rate	-	-	-	1
Penalty parameter C	-	0.0001	-	-

A received waveform from the SS pipe and the location of the 200 best input data points is shown in Figure 4. As with the PMMA pipe these points were taken from the non-saturated part of the waveform. It can be observed that the US waveform decayed much faster in the PMMA pipe when compared to the SS pipe. This is not surprising as acoustic attenuation is much higher in the plastic than steel. In comparison to the PMMA pipe the 200 best input data points for the SS pipe were not all located in one region. For the SS pipe these points were located either between sample numbers 250 to 750 or 1100 to 1600. These results confirmed previous investigations which indicated that when longer waveforms are received, the tail of the wave (waveform points 1000 – 1700 in Figure 4) may be more effected by material at the interface. This is because it may include components of the incident wave which actually propagated into the fouling material (B. Chen et al., 2019).
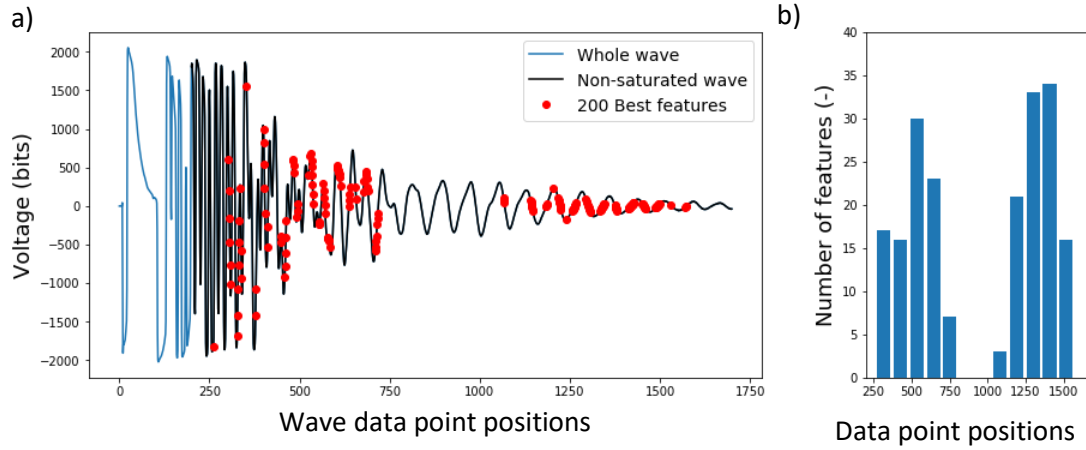
Figure 4: a) Recorded waveform from the SS pipe with location of best input data points. b) Cumulative location of the best input data features.

In general, the optimal number of data points was higher for the SS pipe (Figure 5). This was most likely due to the inclusion of points taken from the proportion of the wave in the tail of the received signal. The only exception to this was the KNN method where a lower number of input data points was optimal. The results in Figure 5 also indicate that the ML model parameters had less impact on the CV scores in the SS pipe than the PMMA pipe. The optimal and selected number of input data points and model parameters can be seen in Table 3 for the SS pipe.
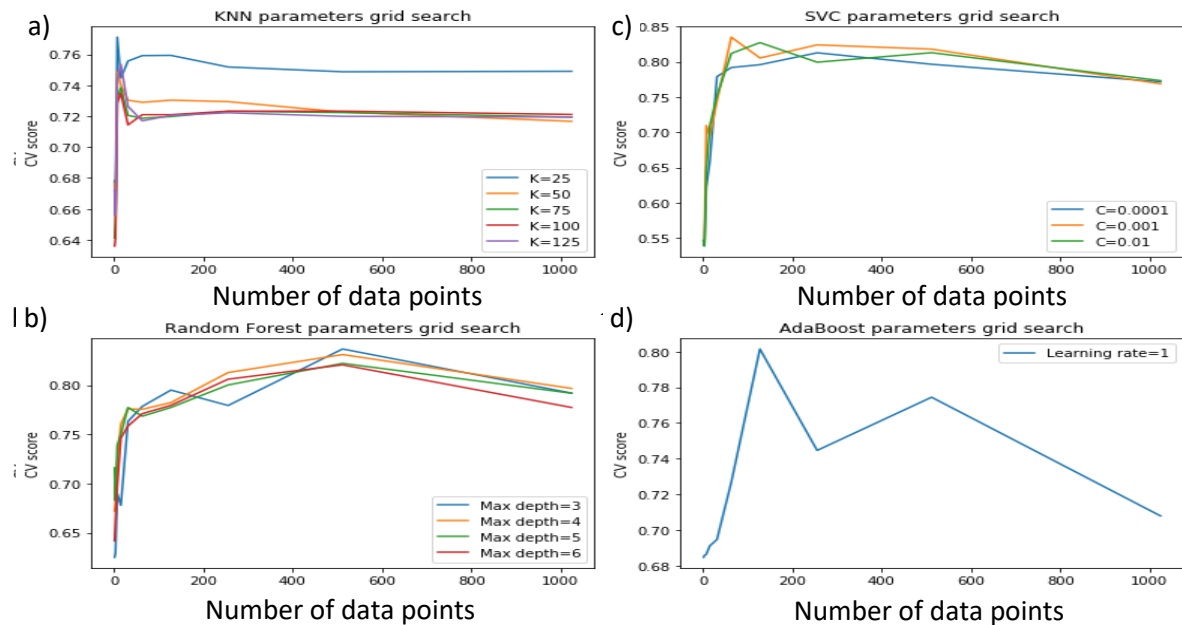

Figure 5: Cross validation (CV) scores for varying number of input data points and other key model parameters for the different modelling methods used in the SS pipe. a) KNN b) SVM c) RF d) AdaBoost.

Table 4: Classification model parameters for PMMA pipe

Model Parameter	KNN	SVM	RF	AdaBoost
Number of input data points	8	64	512	128
Number of neighbours:	25	-	-	-
Number of trees	-	-	500*	500*
Maximum depth	-	-	3	1*
Minimum number of instances in leaves	-	-	10*	10*
Learning rate	-	-	-	1*
Penalty parameter C	-	0.001	-	-

3.2 Classification results from PMMA pipe

The classification models were utilised to make a binary prediction of whether the pipe was clean or dirty. The target data (often called ground truth) was provided by the images. The classification predictions for the tomato fouling at 45℃ and 12℃ are shown in Figure 6, where an orange bar denotes a fouled prediction and a blue bar denotes a clean prediction. The top orange and blue line is the target data taken from the images. These results show 100% classification accuracy for all ML methods for the lower temperature and 100% classification accuracy for all models except the RF for the higher temperature. The incorrect predictions were surprisingly in the first minute of the experiment, but accurate predictions were achieved after this. These results highlight some of the concerns when using methods such as ML, as it is often difficult to determine the cause of incorrect predictions. These could have severe consequence for autonomous systems as they would predict a clean pipe instead of one with fouling resulting in potential hygiene issues or cleaning inefficiencies for a production facility.

a) PMMA Test 1: Tomato 45 °C

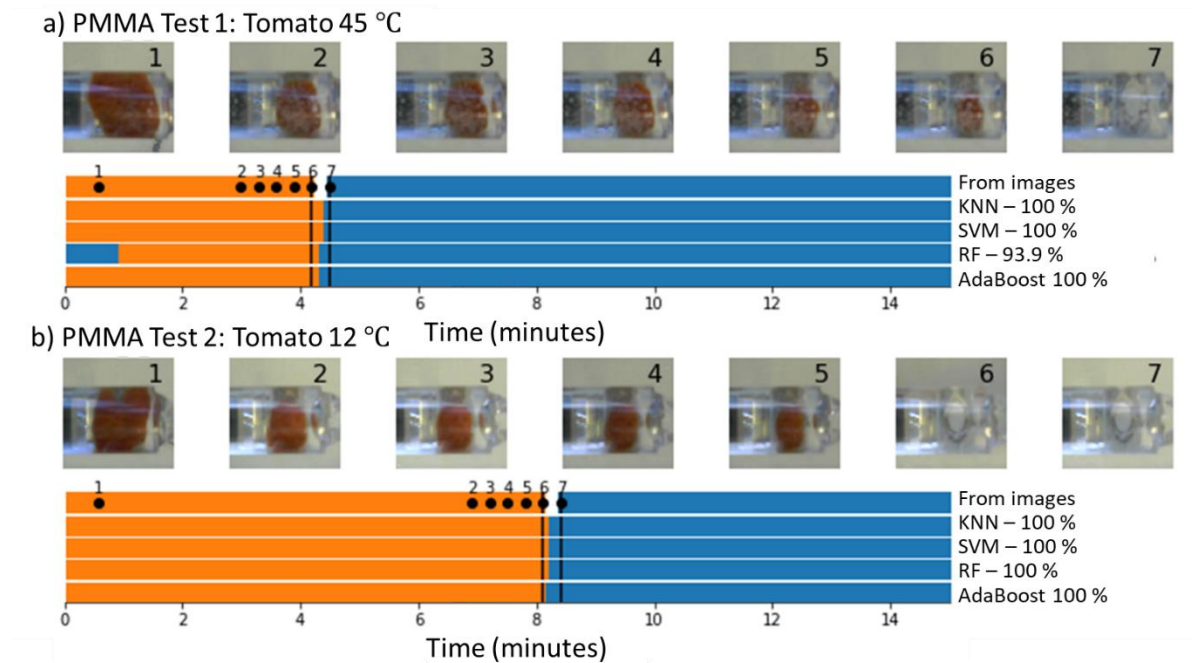b) PMMA Test 2: Tomato 12 °C   Time (minutes)

Figure 6: Classification prediction accuracy for the different machine learning models for tomato in the PMMA pipe. The images of the pipe represent 30 mm in both directions.

The classification predictions for the malt at 45 °C and 12 °C in the PMMA pipe can be seen in Figure 7. In general, the malt classification success rates were lower than the results for tomato (Figure 6). However, they were still high and ranging from 97% to 100%. The only exception to this was the RF model which gave incorrect predictions once the pipe was clean. It was not unexpected to see a difference in results for the different fouling materials studied as previous research has shown that different materials foul and are cleaned from surfaces differently (Fryer & Asteriadou, 2009). It has also been shown in previous research that the removal of different fouling materials also affects ultrasonic measurements (Escrig, Escrig et al., 2019) and the performance of ML classification models (E. Wallhäußer et al., 2013).
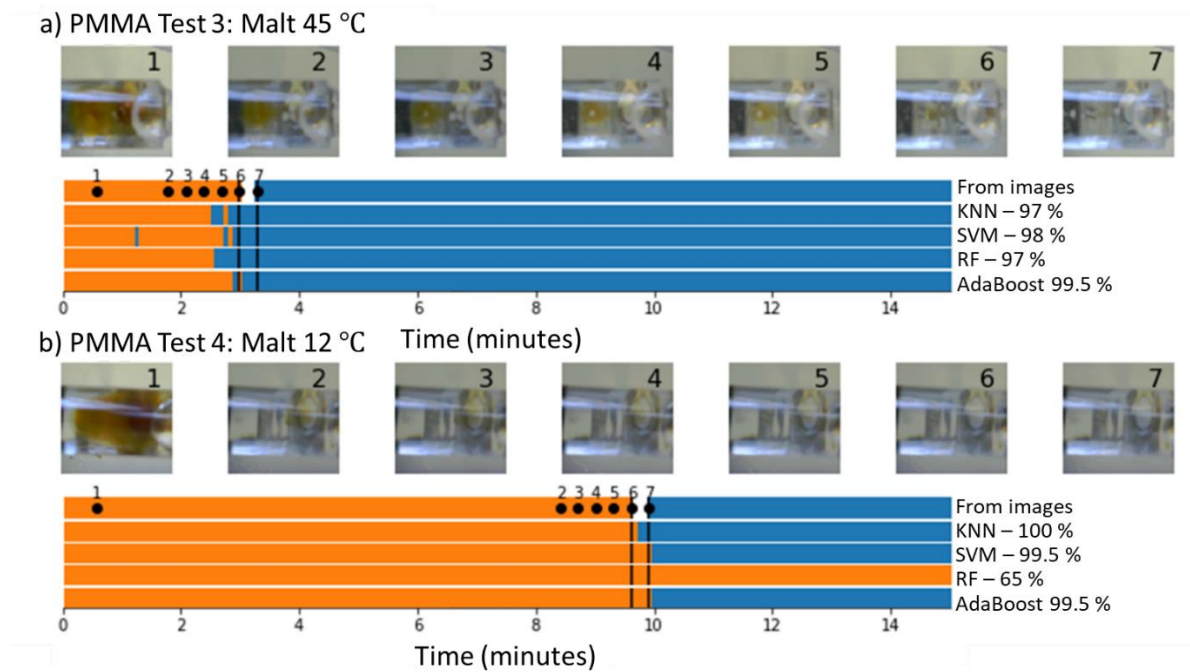
Figure 7: Classification prediction accuracy for the different machine learning models for malt in the PMMA pipe. The images of the pipe represent 30 mm in both directions.

3.3 Classification results from SS pipe

The surface fouling predictions for the tomato (12℃ and 45℃) and malt (12℃ and 45℃) for the SS pipe can be seen in Figure 8. Although camera images were recorded from experiments in the SS pipe, due to light reflections it was not always possible to see when the fouling had been removed. To ensure the data could be labelled correctly visual observations were also performed and the time when the fouling was removed recorded. Due to these difficulties the camera images of the fouling are not shown for the SS pipe in Figure 8. The classification accuracy for all fouling materials and ML methods were between 95% and 100% for the SS pipe. When comparing the two fouling materials there was no major differences, although the models performed marginally better for the Tomato at the higher temperature and the malt at the lower temperature. The incorrect predictions for all materials models and temperatures appeared to be both false positives and false negatives which meant predicting clean when dirty and dirty when clean. The majority of erroneous predictions were located around the point when the fouling was fully removed and the pipe just became clean, which is not surprising as this is the area when the fouling layer would be the smallest. It is also possible that some of the errors were due to how the results were labelled dirty or clean as it was extremely difficult to identify small amounts of fouling on the surfaces. Within a few minutes of the pipe reaching a clean state, the predictions were almost always correct, which would still has significant value in reducing cleaning times.

a) SS Test 1: Tomato 45 ℃

From images
KNN – 99.7 %
SVM – 97.3 %
RF – 99.7 %
AdaBoost 96.6 %

b) SS Test 2: Tomato 12 ℃

From images
KNN – 96.3 %
SVM – 99 %
RF – 95.3 %
AdaBoost 98.3 %

c) SS Test 3: Malt 45 ℃

From images
KNN – 97.3 %
SVM – 93.9 %
RF – 98 %
AdaBoost 92.6 %

d) SS Test 3: Malt 12 ℃

From images
KNN – 97 %
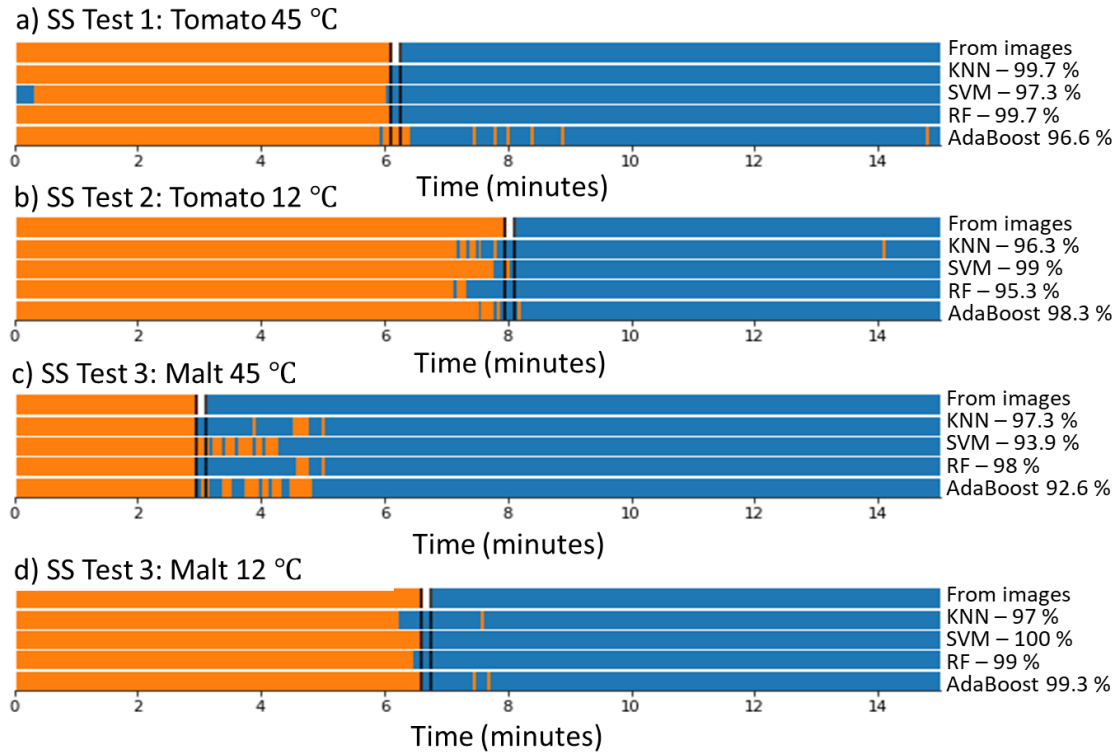SVM – 100 %
RF – 99 %
AdaBoost 99.3 %

Figure 8: Classification prediction accuracy for the different machine learning models for tomato and malt in the SS pipe.

In both pipe materials cleaning experiments were performed at two different temperatures. Although the performance of the classification algorithms were acceptable for both temperatures, in general the classification accuracies were higher for the lower temperature (12℃). The cleaning process was more gradual at 12℃ as dissolution occurs at a slower rate at lower temperature. This gradual cleaning resulted in less classification errors around the time when the fouling was totally removed.

When comparing the results in the SS pipe to the PPMA pipe the main differences where that in the PMMA pipe, 8 of the 16 results had 100% classification accuracy whereas only 1 of the 16 did in the SS pipe. However in the SS pipe the RF models had a similar performance to the other ML methods utilised, which was not the case in the PMMA pipe. These results indicate that the techniques proposed in this work may be more suited to PPMA pipes. However, the classification results were still acceptable for the SS pipes and the reduction in classification accuracy could be attributed to the difficulties in determining when the fouling has been removed in the SS pipe.

All previous work combining US measurements and ML to monitoring surface fouling removal has been on flat test sections. However the results in this work can be compared to these. The previous works showed that classification ML models could successfully be used to predict the presence of surface fouling with prediction accuracies of 98% (Eva Wallhäußer et al., 2011), 100% (E. Wallhäußer et al., 2013), 94% (Eva Wallhäußer et al., 2014) and 98% (Úbeda et al., 2016). It is encouraging that the current results are similar to those reported in the literature, even though the current experiments

were performed in test sections of more complex geometry. Some of these previous works studied different ML methods to determine their effect on classification performance. These works reported that SVM had better prediction accuracies than ANN (E. Wallhäußer et al., 2013) (Eva Wallhäußer et al., 2014). Artificial Neural Networks were not studied in this current work although SVM were and it was found they had similar classification prediction accuracy to the other ML methods studied. Ultrasonic measurements and ML have also been combined to monitor mixing processes (Alexander L Bowler, Bakalis, & Watson, 2020). In this work classification models were developed to predict when mixing was complete and good classification performance was achieved for a range of different ML methods studied. This current and previous works all highlight the benefits of combining US measurements and ML to monitor industrial processes. These works indicate that acceptable classification performance can be achieved for a range of different ML methods although the optimal method for a specific application is not always the same and researchers will often have to adopt a trial and error approach. When developing these models it is always important to consider the significance of the prediction and the level of accuracy required for the specific application.

The ML models studied in this work were either base learners (SVM and KNN) or ensemble methods (RF and AdaBoost). The ensemble methods did not appear to have significantly better prediction accuracy than the base learners but this is most likely because all methods studied performed well. Although this work has shown the potential of US measurements and ML methods as a low cost non-invasive technique for detecting surface fouling in circular pipes and therefore optimising CIP processes, there are some current limitations with the techniques. The primary challenge is labelling the data. In the current method images of the fouling were utilised so that the recorded US waveforms could be labelled dirty or clean. It would not be possible to use a camera in this way in an industrial environment as any equipment would be invasive and have significant difficulties recording usable images. In addition, if a suitable optical system could be used the US measurements system would not be required. To overcome the challenge of labelling data recorded in industrial environments other ML methods should be studied including transfer learning and semi supervised learning. It may also be possible to train particular systems through the use of more invasive methods such as ATP swabbing to determine acceptable clean states.

4 Conclusions

This work has investigated the use of US sensors and ML methods to monitor the removal of surface fouling during cleaning processes. The work has focussed on circular pipes which have not been studied before in the literature. In addition, a range of other variables were studied including pipe material of construction, different food fouling materials and a range of different ML methods. In the work, a clear methodology for data collection and model training is presented, focussing on key

aspects such as determining the optimal number of input data points and other key model parameters using a grid search. The results indicated that the combinations of US measurements and ML models were an effective method for detecting the presence of fouling in circular pipes constructed of either SS or PMMA. The developed models performed well for both fouling materials studied at two different temperatures and prediction accuracy was only slightly better in the PMMA pipes. Of the base and ensemble ML methods utilised no particular technique performed better and there was no clear advantage of using ensemble methods. This work clearly highlights the potential of US measurements and ML models for monitoring and optimising CIP processes and future work should focus on techniques to effectively label data collected in industrial environments, where imaging is not possible.

5   References

Al-Aufi, Y. A., Hewakandamby, B. N., Dimitrakis, G., Holmes, M., Hasan, A., & Watson, N. J. (2019). Thin film thickness measurements in two phase annular flows using ultrasonic pulse echo techniques. *Flow Measurement and Instrumentation*, *66*, 67–78. https://doi.org/10.1016/J.FLOWMEASINST.2019.02.008

Alpaydin, E. (2014). *Introduction to Machine Learning Ethem Alpaydin. Introduction to machine learning*.

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *American Statistician*. https://doi.org/10.1080/00031305.1992.10475879

Berg, T. H. A., Ottosen, N., Berg, F. Van Den, & Ipsen, R. (2017). LWT - Food Science and Technology Inline UV-Vis spectroscopy to monitor and optimize cleaning-in-place (CIP) of whey fi ltration plants. *LWT - Food Science and Technology*, *75*, 164–170. https://doi.org/10.1016/j.lwt.2016.08.014

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (2004). A training algorithm for optimal margin classifiers. https://doi.org/10.1145/130385.130401

Bowler, Alexander L, Bakalis, S., & Watson, N. J. (2020). Monitoring Mixing Processes Using Ultrasonic Sensors and Machine Learning. *Sensors*, *20*(7), 1813. https://doi.org/10.3390/s20071813

Bowler, Alexander Lewis, Bakalis, S., & Watson, N. J. (2019). A review of in-line and on-line measurement techniques to monitor industrial mixing processes. *Chemical Engineering Research and Design*. https://doi.org/https://doi.org/10.1016/j.cherd.2019.10.045

Chen, B., Callens, D., Campistron, P., Moulin, E., Debreyne, P., & Delaplace, G. (2019). Monitoring cleaning cycles of fouled ducts using ultrasonic coda wave interferometry (CWI). *Ultrasonics*, (October 2018), 0–1. https://doi.org/10.1016/j.ultras.2018.12.011

Chen, X. D., Li, D. X. Y., Lin, S. X. Q., & Necati, O. (2003). On-line fouling / cleaning detection by measuring electric resistance — equipment development and application to milk fouling detection and chemical cleaning monitoring. *Science*. https://doi.org/10.1016/S0260-8774(03)00085-2

Cheong, Y. M., Kim, K. M., & Kim, D. J. (2017). High-temperature ultrasonic thickness monitoring for pipe thinning in a flow-accelerated corrosion proof test facility. *Nuclear Engineering and Technology*, *49*(7), 1463–1471. https://doi.org/10.1016/j.net.2017.05.002

Cuckston, G. L., Alam, Z., Goodwin, J., Ward, G., & Wilson, D. I. (2019). Quantifying the e ff ect of solution formulation on the removal of soft solid food deposits from stainless steel substrates.

Journal of Food Engineering, *243*(August 2018), 22–32.
https://doi.org/10.1016/j.jfoodeng.2018.08.018

Eide, M. H., Homleid, J. P., & Mattsson, B. (2003). Life cycle assessment (LCA) of cleaning-in-place processes in dairies. *LWT - Food Science and Technology*, *36*(3), 303–314. https://doi.org/10.1016/S0023-6438(02)00211-6

Escrig, Escrig, J., Woolley, E., Rangappa, S., Simeone, A., & Watson, N. J. (2019). Clean-in-place monitoring of different food fouling materials using ultrasonic measurements. *Food Control*, *104*, 358–366. https://doi.org/10.1016/J.FOODCONT.2019.05.013

Fisher, O., Watson, N., Porcu, L., Bacon, D., Rigley, M., & Gomes, R. L. (2018). Cloud manufacturing as a sustainable process manufacturing route. *Journal of Manufacturing Systems*, *47*. https://doi.org/10.1016/j.jmsy.2018.03.005

Fratamico, P. M., Annous, B. A., & Guenther, N. W. (2009). Biofilms in the food and beverage industries. *Biofilms in the Food and Beverage Industries*, (December), 1–580. https://doi.org/10.1533/9781845697167

Fryer, P. J., & Asteriadou, K. (2009). A prototype cleaning map : A classification of industrial cleaning processes. *Trends in Food Science & Technology*, *20*(6–7), 255–262. https://doi.org/10.1016/j.tifs.2009.03.005

Fryer, P. J., Christian, G. K., & Liu, W. (2006). How hygiene happens: Physics and chemistry of cleaning. *International Journal of Dairy Technology*, *59*(2), 76–84. https://doi.org/10.1111/j.1471-0307.2006.00249.x

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Guérin, R., Ronse, G., Bouvier, L., Debreyne, P., & Delaplace, G. (2007). Structure and rate of growth of whey protein deposit from in situ electrical conductivity during fouling in a plate heat exchanger. *Chemical Engineering Science*, *62*(7), 1948–1957. https://doi.org/10.1016/j.ces.2006.12.038

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. *Bayesian Forecasting and Dynamic Models* (Vol. 1). https://doi.org/10.1007/b94608

Lohr, K. R., & Rose, J. L. (2003). Ultrasonic guided wave and acoustic impact methods for pipe fouling detection. *Journal of Food Engineering*, *56*(4), 315–324. https://doi.org/10.1016/S0260-8774(02)00156-5

Lyndgaard, C. B., Rasmussen, M. A., Engelsen, S. B., Thaysen, D., & Van Den Berg, F. (2014). Moving from recipe-driven to measurement-based cleaning procedures: Monitoring the Cleaning-In-Place process of whey filtration units by ultraviolet spectroscopy and chemometrics. *Journal of Food Engineering*, *126*, 82–88. https://doi.org/10.1016/j.jfoodeng.2013.10.037

Pereira, A., Mendes, J., & Melo, L. F. (2009). Monitoring cleaning-in-place of shampoo films using nanovibration technology. *Sensors and Actuators, B: Chemical*, *136*(2), 376–382. https://doi.org/10.1016/j.snb.2008.11.043

Pettigrew, L., Blomenhofer, V., Hubert, S., Groß, F., & Delgado, A. (2015). Optimisation of water usage in a brewery clean-in-place system using reference nets. *Journal of Cleaner Production*, *87*, 583–593. https://doi.org/10.1016/j.jclepro.2014.10.072

Polat, K., & Güneş, S. (2009). A new feature selection method on classification of medical datasets: Kernel F-score feature selection. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2009.01.041

Ruppert, D. (2004). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *Journal of the American Statistical Association*. https://doi.org/10.1198/jasa.2004.s339

Sabzevari, M., Martínez-Muñoz, G., & Suárez, A. (2018). Vote-boosting ensembles. *Pattern Recognition*. https://doi.org/10.1016/j.patcog.2018.05.022

Simeone, A., Deng, B., Watson, N., & Woolley, E. (2018). Enhanced Clean-In-Place Monitoring Using Ultraviolet Induced Fluorescence and Neural Networks, *c*. https://doi.org/10.3390/s18113742

Simeone, A., Watson, N., Sterritt, I., & Woolley, E. (2016). A Multi-sensor Approach for Fouling Level Assessment in Clean-in-place Processes. *Procedia CIRP*, *55*, 134–139.

https://doi.org/10.1016/J.PROCIR.2016.07.023

Tamachkiarow, A., & Flemming, H. C. (2003). On-line monitoring of biofilm formation in a brewery water pipeline system with a fibre optical device. *Water Science and Technology*, *47*(5), 19–24. https://doi.org/10.2166/wst.2003.0270

Tlili, M. M., Rousseau, P., Ben Amor, M., & Gabrielli, C. (2008). An electrochemical method to study scaling by calcium sulphate of a heat transfer surface. *Chemical Engineering Science*, *63*(3), 559–566. https://doi.org/10.1016/j.ces.2007.09.035

Úbeda, M. A., Hussein, W. B., Hussein, M. A., Hinrichs, J., & Becker, T. M. (2016). Acoustic sensing and signal processing techniques for monitoring milk fouling cleaning operations. *Engineering in Life Sciences*, *16*(1), 67–77. https://doi.org/10.1002/elsc.201400235

Van Asselt, A. J., Van Houwelingen, G., & Te Giffel, M. C. (2002). Monitoring system for improving cleaning efficiency of cleaning-in-place processes in dairy environments. *Food and Bioproducts Processing: Transactions of the Institution of of Chemical Engineers, Part C*, *80*(4), 276–280. https://doi.org/10.1205/096030802321154772

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*. https://doi.org/10.1109/72.788640

Wallhäußer, E., Hussein, M. A., & Becker, T. (2012). Detection methods of fouling in heat exchangers in the food industry. *Food Control*, *27*(1), 1–10. https://doi.org/10.1016/j.foodcont.2012.02.033

Wallhäußer, E., Hussein, W. B., Hussein, M. A., Hinrichs, J., & Becker, T. (2013). Detection of dairy fouling: Combining ultrasonic measurements and classification methods. *Engineering in Life Sciences*, *13*(3), 292–301. https://doi.org/10.1002/elsc.201200081

Wallhaußer, E., Hussein, W. B., Hussein, M. A., Hinrichs, J., & Becker, T. M. (2011). On the usage of acoustic properties combined with an artificial neural network - A new approach of determining precense of dairy fouling. *J. Food Eng.*

Wallhäußer, Eva, Hussein, W. B., Hussein, M. A., Hinrichs, J., & Becker, T. M. (2011). On the usage of acoustic properties combined with an artificial neural network - A new approach of determining presence of dairy fouling. *Journal of Food Engineering*, *103*(4), 449–456. https://doi.org/10.1016/j.jfoodeng.2010.11.015

Wallhäußer, Eva, Sayed, A., Nöbel, S., Hussein, M. A., Hinrichs, J., & Becker, T. (2014). Determination of cleaning end of dairy protein fouling using an online system combining ultrasonic and classification methods, 506–515. https://doi.org/10.1007/s11947-012-1041-0

Wang, W., Liu, X., Li, X., Xu, G., & Zhang, S. (2019). Modeling Flaw Pulse-Echo Signals in Cylindrical Components Using an Ultrasonic Line-Focused Transducer with Consideration of Wave Mode Conversion. *Sensors (Basel, Switzerland)*, *19*(12). https://doi.org/10.3390/s19122744

Watson, N. J. (2015). Ultrasound Tomograghy. In M. Wang (Ed.), *Industrial Tomography: systems and applications* (pp. 235–262). Woodhead Publishing.

Withers, P. (1994). Ultrasonic sensor for the detection of fouling in UHT processing plants. *Food Control*, *5*(2), 67–72. https://doi.org/10.1016/0956-7135(94)90088-4

Withers, P. M. (1996). Ultrasonic, acoustic and optical techniques for the non-invasive detection of fouling in food processing equipment. *Trends in Food Science and Technology*, *7*(9), 293–298. https://doi.org/10.1016/0924-2244(96)10031-5