



# Use of acoustic emission in combination with machine learning: monitoring of gas–liquid mixing in stirred tanks

Giuseppe Forte<sup>1,2</sup> · Federico Alberini<sup>2</sup> · Mark Simmons<sup>2</sup> · Hugh E. Stitt<sup>1</sup>

Received: 19 July 2019 / Accepted: 22 June 2020 / Published online: 2 July 2020  
© The Author(s) 2020

## Abstract

Operations involving gas–liquid agitated vessels are common in the biochemical and chemical industry; ensuring good contact between the two phases is essential to process performance. In this work, a methodology to compute acoustic emission data, recorded using a piezoelectric sensor, to evaluate the gas–liquid mixing regime within gas–liquid and gas–solid–liquid mixtures was developed. The system was a 3L stirred tank equipped with a Rushton Turbine and a ring sparger. Whilst moving up through the vessel, gas bubbles collapse, break or coalesce generating sound waves transmitted through the wall to the acoustic transmitter. The system was operated in different flow regimes (non-gassed condition, loaded and complete dispersion) achieved by varying impeller speed and gas flow rate, with the objective to feed machine learning algorithms with the acoustic spectrum to univocally identify the different conditions. The developed method allowed to successfully recognise the operating regime with an accuracy higher than 90% both in absence and presence of suspended particles.

**Keywords** Acoustic emission · Gas–liquid mixing · Stirred tank · Machine learning

## Introduction

Improving process monitoring is a common need within the process industry, including chemical, food, biochemical and pharmaceuticals (Boyd and Varley 2001). To further advance the field, innovation is needed not only in the theoretical foundations but also in identifying new technical solutions (Hu et al. 2015). Desirable features for process measurements are to be non-invasive and suitable for in-line application with real time response, in order to avoid delays in intervening with control measures. Amongst several potential sensing methods, acoustic emission (AE) is a low cost, data-rich technique with applicability for in-line monitoring. Traditionally acoustic techniques are categorised as either with active or passive acoustics. The former consists of a transmitter generating an acoustic wave within the system and a receiver acquiring the response of the stimulated system. The latter, also known as acoustic emission (AE), is

composed only by a sensor recording acoustic waves generated by the process itself.

The signal will often be a combination of numerous acoustic events all propagating to the sensor via different paths. Most of AE research has focused on fault detection (leakage, failure,...) (Boyd and Varley 2001), corrosion (Cole and Watson 2005), grinding (Griffin and Chen 2016) and tool wearing (Elforjani and Shanbr 2018; Li et al. 2015). More recent studies have also investigated AE as a means of monitoring physico-chemical changes within processes involving powder and fluids. For example, Aldrich et al. (Aldrich and Theron 2000) have used directional microphones to estimate particle size in a ball mill using continuous regression. Other studies have applied AE to powder pneumatic conveying (Esbensen et al. 1998) and V-blenders (Crouter and Briens 2015). Applications in multiphase mixing are also reported in the literature, Nordon et al. (2004), for example, applied AE to a jacketed stirred tank for monitoring of a heterogeneous reaction. While the first known reference on gas–liquid flow dates from the 1920s (Bragg Sir 1921), more recent works (Addali et al. 2010) have used AE energy information to predict gas phase fraction in a two phase (air–water) slug flow.

Although, many researchers have investigated various methods to determine multiphase mixing regimes in stirred

✉ Federico Alberini  
f.alberini@bham.ac.uk

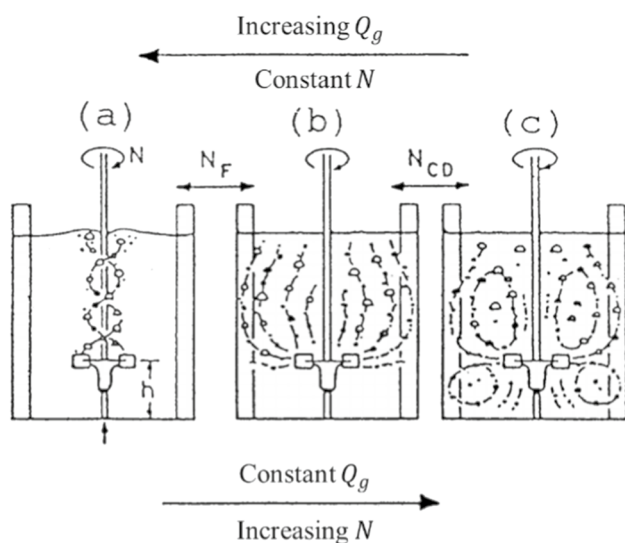
<sup>1</sup> Johnson Matthey Technology Centre, Billingham TS23 4LB, UK

<sup>2</sup> School of Chemical Engineering, University of Birmingham, Edgbaston campus, Birmingham B15 2TT, UK

tanks in real time, as reported in subsequent sections, their applications have been limited to R&D. This is mainly due to difficulties in retrofitting the devices to existing plants as well as to slow responses. In this work, AE is applied to monitor gas–liquid mixing in a 3L stirred tank, with the objective to identify the operating bubble dispersion regime. This work aims to propose a methodology to obtain real time information by installing an acoustic piezoelectric sensor on the outside of a stirred tank. The data-richness and high time resolution of the used technique represent desirable features for modern control system strategies.

## Gas–liquid mixing

Gas–liquid reactors are very common unit operations in the biochemical and chemical industries, where ensuring appropriate interphase contact and gas dispersion is critical. Different regimes are observed when sparging gas into stirred tanks; namely flooded, loaded, completely dispersed, and gas recirculation (Nienow et al. 1985) depending on gas flow rate, tank design and fluids properties (e.g. rheology, density, interphase tension). The different regimes (shown in Fig. 1) characterise how well the gas is sparged within the volume of the tank: contact between the two phases increases moving from the flooded (a) towards the completely dispersed condition (c). In the flooding regime the impeller speed  $N$  does not overcome the flooding critical speed  $N_f$ , therefore the gas dispersion is minimum, where a plume of gas rises from the sparger up to the surface along the impeller. At higher impeller speeds when the  $N$  is higher than  $N_f$ , but lower than  $N_{cd}$  (the completely dispersed critical speed), the system is in the loading regime; where the gas is more



**Fig. 1** Pictorial representation of the main gas–liquid regime: **a** flooded, **b** loaded and **c** completely dispersed (Nienow et al. 1985)

dispersed but does not cover the whole volume of the tank as the radial drag force is overcome by the buoyancy forces. At higher impeller speed ( $N_{cd} < N < N_r$ ), in the completely dispersed regime, the gas is well dispersed through the whole vessel at low power, but it is beyond  $N_r$  (the recirculation critical speed) that the full gas recirculation regime is observed. The latter two are generally the desirable conditions for unit operations requiring good contact between the two phases.

Commonly, the operating regime is estimated based on a characteristic regime chart (Middleton 1992) based on the definition of two dimensionless numbers: the Froude number ( $Fr = N^2 D/g$ ), defined as the ratio of the flow inertia to the gravity field and the gas flow number (Warmoeskerken and Smith 1985) ( $Fl = Q_g/(ND^3)$ ) representing the ratio of the gas inlet rate to the impeller pumping.

## Sensing and measurement techniques

Existing measurement techniques for identifying the different regimes are multiple, including optical (Yawalkar et al. 2002), level probes (Gao et al. 2001), ultrasonics (Cents et al. 2005) and tomography [Electrical Resistance ERT (Forte et al. 2019; Jamshed et al. 2018), X-ray (Ford et al. 2008),  $\gamma$ -ray (Veera et al. 2001)], but are often limited by practical considerations (invasiveness, retrofitting, opacity, safety) and cost issues. The mentioned systems all require high investment cost in terms of devices and operations. They are also usually fixed solutions characterized by rather heavy retrofitting. Furthermore, in cases of high energy based methods (X-ray,  $\gamma$ -ray) an additional cost is represented by the safety measures needed to ensure safe use of the technique. The AE device is, instead, composed by a simple hardware which does not require retrofitting and is also a portable solution. Indeed, the sensor can be placed on a tank and moved to another quickly, being installed on the outside wall. This, together with the low investment and operational cost make it convenient to use and furthermore devoid of safety issues. The AE system applied in this study has, therefore, the potential to be a cheap, easy to install (and retrofit) alternative technique that can offer reliable information on the operating gas–liquid regime.

In this work, AE emission data are acquired at loaded and completely dispersed regimes as well as in ungasged condition, to simulate accidental shut down of air feed. Machine learning (ML) supervised algorithms are used to identify and recognize the different conditions. To evaluate the robustness of the technique in correctly identifying the operating regime, a deviation from the biphasic condition is also introduced, by adding to the bulk solid particles up to a concentration of 5% w/w. This occurrence is not unusual in case of chemical reactions, as for example solid precipitation, occurring from mixing reactive gas–liquid mixtures

in stirred tanks (Zhao et al. 2016, 2017). The presence of solids will cause an increase in the emitted signal because of impacts of the particles with the wall, the impeller and with each other, but the system will be challenged to correctly predict the gas–liquid operating regime. The regime is considered unchanged at fixed air volumetric flow and impeller speed: for low concentration [ $< 20\%$  w/w (Chapman et al. 1983b)] of solids, works in the literature report little effects of the presence of the particles on the gas sparging dynamics (Bao et al. 2006; Chapman et al. 1983b).

## Machine learning algorithms

The use of machine learning algorithms in AE applications has recently and successfully been applied by other researchers in the literature for condition monitoring of cutting tools (Chen et al. 2011; Li et al. 2015), gear box fault diagnosis (Li et al. 2016) and for data clustering (Pomponi and Vinogradov 2013). The implementation of such techniques in chemical and manufacturing industry has recently seen increasing interest since such data-driven and statistical approaches fit well with the smart manufacturing/Industry 4.0 trend (Wuest et al. 2016).

Techniques involving Artificial Intelligence (AI) have been applied in several fields including, game playing, robotics, facial and speech recognition (Venkatasubramanian 2019) with the first attempt to chemical engineering applications reported in the 80s for catalyst design (Banares-Alcantara et al. 1985; Bañares-Alcántara et al. 1987), application that is still object of current studies (Ulissi et al. 2017). Recent applications in researches of industrial interest can be found in estimation of physical properties of organic molecules (National Academies of Sciences 2018), shape memory alloys design (Xue et al. 2016) and studies on colloidal self-assembly systems (Spellings and Glotzer 2018). The spreading of such investigations have in the last years interested a much wider range of operations thanks to cheaper and faster calculators and fast access to large memory storage (cloud) (Venkatasubramanian 2019). Furthermore, also the psychological barriers within organization, both within management and manufacturing have lowered thanks to spreading of AI devices within domestic and personal application (Amazon Alexa is an example).

Many ML techniques including both unsupervised and supervised algorithms have been developed. In this study, some of the more common machine learning methods are used for processing the AE data in the frequency domain, to recognize the different operative conditions; solving a classification problem (Wu et al. 2008). The herein used algorithms are: logistic regression, support vector machine (SVM), k-nearest neighbour (k-NN) and decision tree; of which the last three were implemented in MATLAB®

classification learner application within the Statistic and Machine Learning toolbox.

Logistic regression has already been applied in processing acoustic emission data for predicting reliability of cutting tools and for condition monitoring of bearing elements (Rozak et al. 2018) to provide adequate maintenance schedules (Li et al. 2015). It is a nonlinear statistical method which has been extensively used also beyond machine reliability and life prediction (Caesarendra et al. 2010; Yan et al. 2004; Yan and Lee 2004) in the field of economics (Martin 1977) and health (Bender and Kuss 2010). The algorithm itself can be defined as a binomial regression (Hilbe 2009), where the output is a probability  $h_\theta$  of a condition being verified, in the specific case the belonging of the signal to one specific class. Logistic regression makes use of a logistic function, also known as sigmoid function:

$$h_{\theta^i}^i(x) = 1 / (1 + e^{-\theta^i x}) \quad (1)$$

In the equation above  $h_{\theta^i}^i(x)$  is the obtained probability, that the input variable  $x$  (the AE spectrum in this study) belongs to a given class  $i$  (one of the three operating conditions).  $\theta^i$  is the vector of parameters (frequency features in the spectrum), that the machine tunes in the learning step. For each condition  $i$  a parameter  $\theta^i$  is obtained in the learning process. Such process consists of an optimisation on the parameter  $\theta^i$  run by the machine, using the gradient descent strategy, to correctly classify the training data set. The test, instead, sees the machine providing the probability that the fed spectrum belongs to the three classes and it will assign it to the one with highest probability  $h_{\theta^i}(x)$ .

The decision tree algorithm (Murthy 1998) has a node structure for addressing the classification problem using hierarchical, sequential binary classifications. The tree structure is built with an increasing number of nodes moving from the top of the tree towards the bottom branches. The nodes composing the tree are ranked during the learning process where the features flagged as highly informative are positioned at the top nodes of the tree; moving towards the bottom level features, where less informative scores are placed. This algorithm is widely used because of its transparent mechanism that help users to visualise how the classification algorithm takes its decision (Jang 1993). Among different decision tree algorithms (Kotsiantis 2007), one of the most frequently used due to its effectiveness and simplicity, is the Iterative Dichotomiser 3 (ID3) (Soofi and Awan 2017). This algorithm is described by a parameter called maximum number of splits, which is a control parameter of the final depth of the tree (Safavian and Landgrebe 1991): the higher this number, the larger, more detailed and complex will be the tree. Within MATLAB® Classification Learner toolbox, two decision tree algorithms were selected for test based on

this control parameter: fine tree (100 maximum number of splits) and medium (20 maximum number of splits).

A different classification logic characterises the  $k$ -Nearest-Neighbour ( $k$ -NN) algorithms (Cover and Hart 1967). Each instance is assigned to a certain class based on the class to which its nearest neighbours belong. The principle is that, in the vector space defined by the number of used features, instances within a dataset will be closer to instances belonging to the same class. The definition of distance between data points was computed using the Euclidean definition (Wu and Zhang 2002), although many other methods can be used (Canberra, Chebyshev, Minkowsky etc.) (Prasath et al. 2017). One of the critical parameters defining the algorithm is  $k$  which describes the number of nearest neighbours that the algorithm considers in assigning each instance to a class. In this work, three values of  $k$  are used for the analysis: 1 (termed as fine  $k$ -NN), 10 (medium  $k$ -NN) and 100 (coarse  $k$ -NN).  $k$ -NN algorithms are widely used in applications such as face recognition (Kasemsumran et al. 2016), traffic forecasting (Zhang et al. 2013) and speaking recognition (Rizwan and Anderson 2014) achieving high performance when large training datasets are available. It is robust to noisy data and easy to visualize but it often requires large memory allocation (Soofi and Awan 2017).

Support vector machine (SVM) is considered one of the most accurate and robust algorithms among the common methods (Vapnik 2000). It is an efficient and quick method especially in the training step. Given a dataset, SVM finds the best classification function that divides the instances in two classes. The “best” function is identified geometrically and can be a hyperplane in the case of linear classification or it can have different shapes depending on the definition of the kernel function that characterises the method. Amongst others, this function may be linear, parabolic, hyperbolic. Defined the kernel function, the SVM finds the parameters within this function that maximise the margin between the classes (Suykens and Vandewalle 1999). The margin

corresponds to the shortest distance between the closest points to the boundary hyperplane function. SVM is able to deal with a large variety of classification problems including high dimensional and non-linear problems (Soofi and Awan 2017). Although SVM is very powerful, it is difficult to visualise (Karamizadeh et al. 2014) and require accurate often a priori selection of a number of parameters within the kernel function.

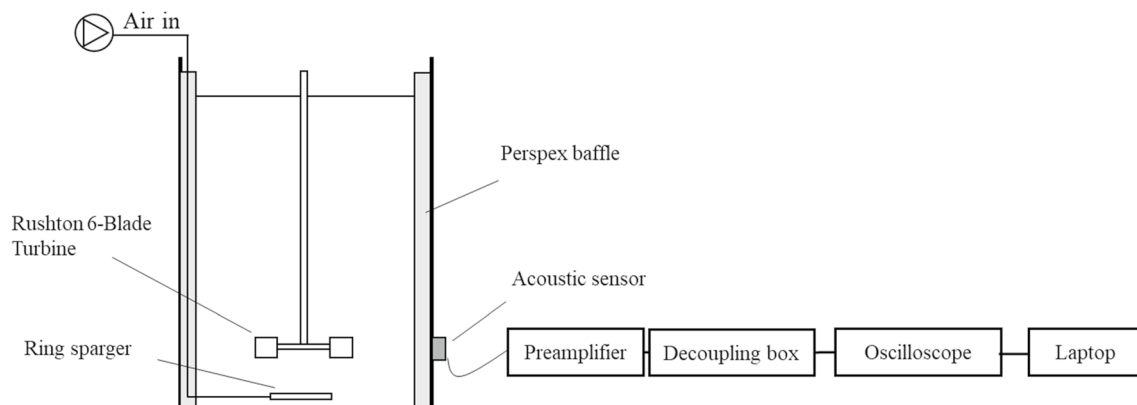
Among the wide spectra of algorithms and kernels available in the literature, in this work some of the most established methods are used to process data-rich AE signal in order to identify the operating gas–liquid regime in two-phases and three-phases conditions. The objective is to develop a flexible method that can be extended to different systems and evaluate expected accuracy using a supervised machine learning approach.

## Materials and methods

### Stirred tank configuration

The agitated system, already used by Forte et al. (2019), was a cylindrical Perspex tank equipped with a gas sparger (ring shape) at the bottom of the tank. The vessel, having a diameter,  $T$ , of 0.14 m, was equipped with four Perspex baffles with width,  $B$ , equal to  $T/10$ . A stainless steel six blade Rushton Disc Turbine (RDT6) with diameter  $D=0.056$  m ( $D/T=2/5$ ) was used for stirring. The used liquid was an aqueous solution of Nickel nitrate hexahydrate (99.99% Sigma Aldrich®) and its level was set to 0.21 m ( $H/T=3/2$ ). In Fig. 2 a schematic of the tank is reported, together with the AE measurement device scheme.

The air flow (ranging between 5 and 10 L min<sup>-1</sup>, corresponding to 1.5 and 3 vvm respectively) was fed from a ring sparger positioned at the bottom of the tank and equipped with 8 orifices of 0.5 mm. During the different experiments,



**Fig. 2** Schematic of the stirred tank and the AE equipment

the Reynolds number was kept between 15,000 and 60,000 in order to have, in all cases, turbulent regime. The impeller speed (300–1300 rpm) and the gas flow rate were changed to set one of the operating regimes (loading, complete dispersion) and then the same conditions of impeller speed were recorded with no air fed to the system (to mimic failure of gas feed unit). The achievement of the different regimes was theorised using flow regime maps and visually checked throughout the experiments, [using optical methods as per (Forte et al. 2019)]. In Table 1, operating conditions for the different regimes are reported.

In the three-phase experiment, stainless steel (AISI 316) spherical particles (Alpha Aesar ©) were used as solid phase. The particles, used at concentrations of 3, 4 and 5% w/w had a density  $\rho_s$  of approximately  $8000 \text{ kg m}^{-3}$  and size between 0.177 and 0.420 mm. The same operating points as for Table 1 were taken.

### Acoustic emission

The used apparatus has been assembled as for similar works in the literature (Nordon et al. 2004): a piezoelectric sensor (Vallen Systeme GmbH, Icking, Germany) with resonance frequency of 375 kHz and diameter 20.3 mm, was attached to the tank using a silicone based vacuum grease to ensure acoustic coupling with the tank. A preamplifier (40 dB gain, Vallen Systeme GmbH) was also part of the measurement rig as well as a decoupling box (Vallen Systeme GmbH) that removed the electrical noise introduced by the preamplifier before feeding the signal to the oscilloscope (5243A Pico®

Technology Limited), used to record data, arranged as for Fig. 1. The sensor was placed on the vessel outside wall at a height of 0.04 m from the bottom of the tank, corresponding to the impeller region. Twenty-five measurements at each combination of impeller speed and gas flow rate were taken for 0.2 s with a sampling rate of 1000 kHz, over a maximum recording capability of 20 MHz.

### Data processing and machine learning

The purpose of the present work is to develop a methodology to interrogate the information from the acoustic spectrum to determine the operating regime based on a previously trained model. The acquired signals were pre-processed before they were fed to the machine learning algorithm. A schematic summary of the data processing is shown in Fig. 3.

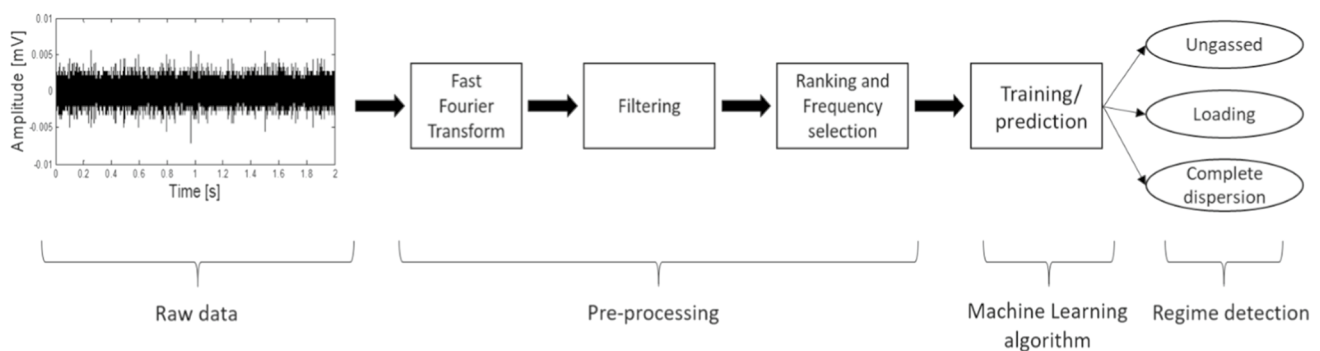
The time-domain signals were processed using the Fast Fourier Transform function (Cooley et al. 1969) to obtain the corresponding dataset in the frequency domain. Examples of time domain and frequency domain signals for loading and complete dispersion conditions are reported in Fig. 4. Each of the acquired acoustic spectra is composed by over 200,000 points, hence frequencies.

A feature scaling and mean normalisation were then applied to the obtained spectra, in order to have values at each frequency ranging from  $-1$  to  $1$ , to avoid biasing the processing towards the features exhibiting the highest amplitude (Kouroussis et al. 2000).

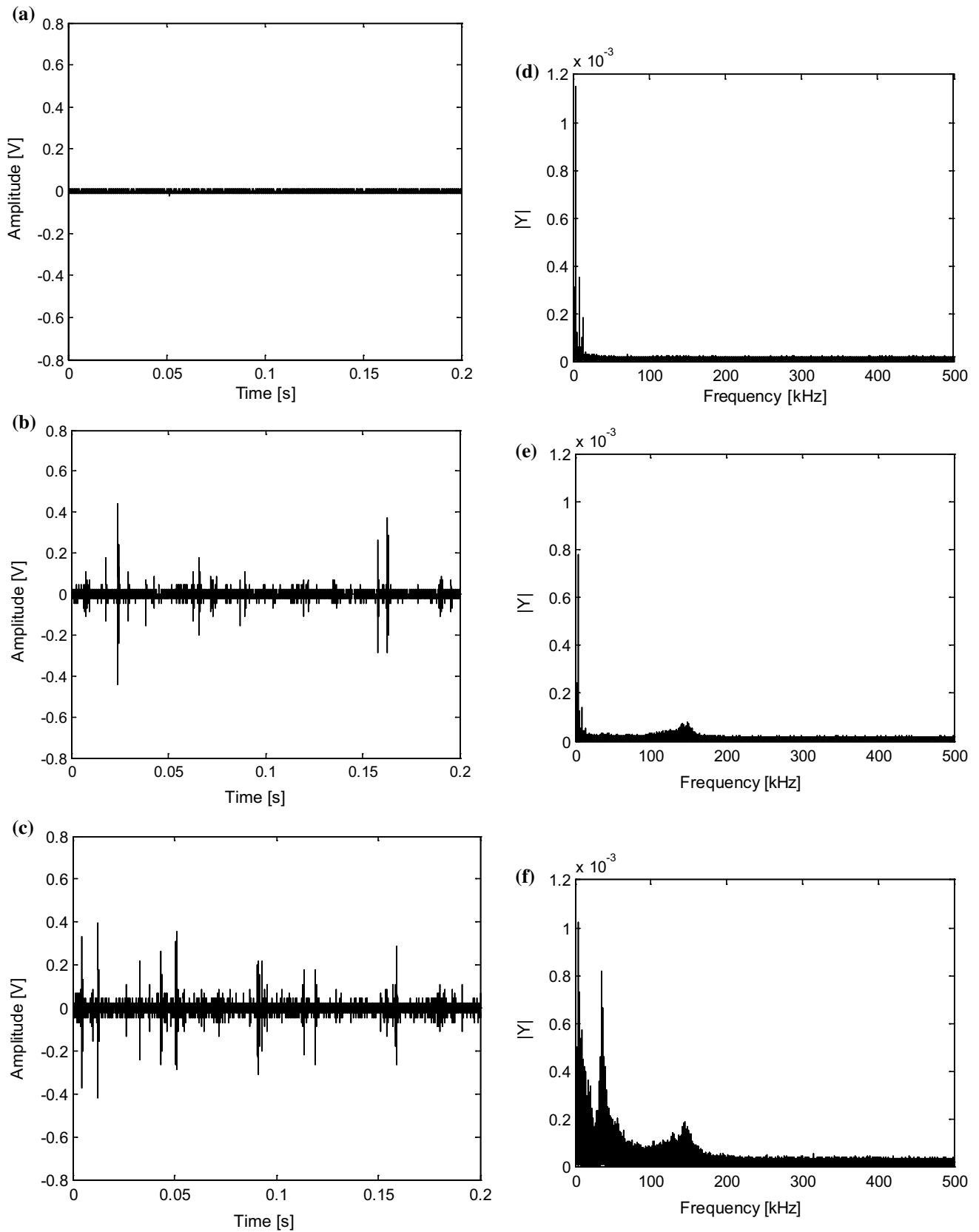
Low frequencies are notoriously prone to noise propagation and environmental interference (Nordon et al. 2006;

**Table 1** Operating conditions for the different regimes

Regime	Fr	Fl	Impeller speed (rpm)	Air flow rate (vvm)
Ungassed	0.24–2.55	–	400–1300	–
Loading	0.24–0.55	0.08–0.0529	400–600	1.5–3
Complete dispersion	1.0–2.55	0.0244–0.08	800–1300	1.5–3



**Fig. 3** Flow diagram summarising the data processing of AE signal



**Fig. 4** Time domain (a, b, c) and frequency domain (d, e, f) signals for ungasged (a, d), loaded (b, e) and complete dispersion conditions (c, f)



Whitaker et al. 2000), therefore frequencies below 4 kHz were removed from the analysis.

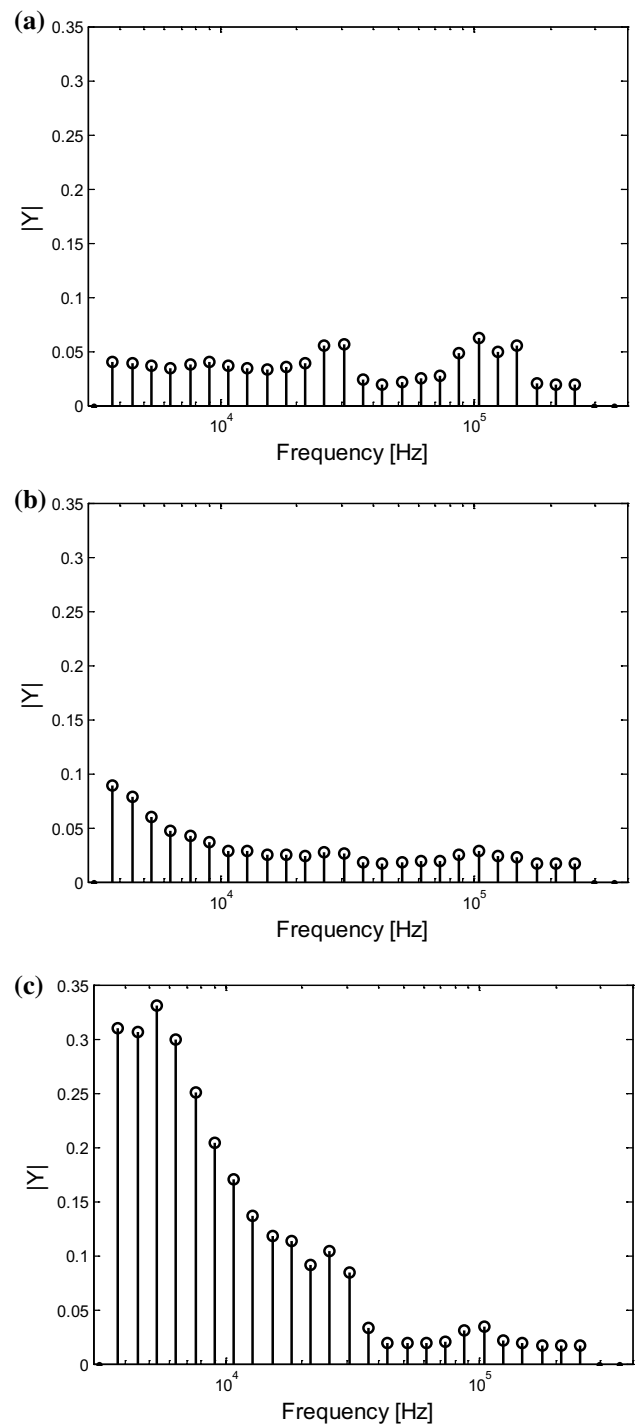
Implementation of machine learning algorithms requires a high number of computational operations. At the same time, high number of points may be cause of bias in the implementation of the classification algorithm. In order to reduce the interrogated features, a method based on the highest variance in the data is proposed. The frequencies in the spectrum are ranked based on their variability across the data, with the most variable ranked as 1st and so on. Based on this ranking, a number  $n$  of the most variable frequency peaks was selected as features characterising each instance before being fed to the machine. This is done to reduce the amount of processed data, as the objective of the study is to investigate AE suitability for inline real-time installation. The selection of  $n$  number of used features is investigated with  $n_{max}=30,000$  being the maximum value. The  $n_{max}$  value is the number for which the cumulative variance is 99.9% of the observed variability across the spectra. In Fig. 5, an average frequency domain plot is reported for the three conditions for  $n=n_{max}$  to show the span of used frequencies across the whole spectrum. In Fig. 5 the 30,000 frequencies are grouped in logarithmic scale and plotted using a Stem Plot, for a clearer visual representation of the large number of data.

As seen in Fig. 5, the selected frequencies in the spectra for the three conditions neglect the frequencies over 200 kHz. It must be noted that the logarithmic grouping will use a wider range for the higher frequencies, hence, the most variable features in the case of  $n=n_{max}$  are for frequencies in the audible range (< 20 kHz). This is more evident by increasing the dispersion of gas at the complete dispersion regime: the peaks at lower frequencies increase their intensity.

Following the feature selection step, the data are divided in three datasets for the assessment:

- A training dataset, consisting in 60% of the acquired data, was fed to the machine for the training process, together with their own corresponding class of belonging.
- A cross-validation dataset, consisting in 20% of the acquired data, used for selecting the optimum number of used features.
- A test dataset, consisting in the remaining 20% of the data, unseen by the machine was used to evaluate the final accuracy of the method.

The three dataset are built taking for each operating point (fixed impeller speed and gas feed rate) the corresponding portion in percentage (Example: for a given operating point 25 AE spectrum are recorded, of which 15 are included in the training dataset, 5 in the cross-validation dataset and the remaining 5 in the test dataset).



**Fig. 5** Frequency distribution across the spectrum for ungasged (a), loaded (b) and complete dispersion (c) regime

The task is addressed as a classification problem: in the training step, the system gets as an input the spectra and the corresponding class (ungassed, loaded, completely dispersed). Depending on the used algorithm, the machine builds criteria to identify the correspondence between the AE data and their known operating condition. When the

training is complete, given a new AE set of data the system will provide as an output the predicted operating condition.

The training process is carried out by obtaining acoustic signal at different impeller speed and air flow rate at the three regimes. The cross-validation step consists of identifying the optimum number of features ( $n$ ) to use for the analysis. The criteria to automatically select  $n$  are dependent on two selected parameters:

- the processing time; in this case identified as the time taken by the CPU to process the learning step, arbitrarily decided to be lower than 1 s;
- the average accuracy of the methods; four variants are chosen from each family of the used algorithms and their accuracy is averaged for identifying the zone of optimum.

In the test step, the performance of the algorithms are calculated using two factors, one being the accuracy of the method, calculated as the ratio between number of cases correctly predicted and the total cases number; the second factor is known as the F1-score (Powers 2011). It is widely used in medical applications for diagnosis methods (Ferizi et al. 2019; Rink et al. 2011) and it takes into account positive prediction as well as false negative. It is defined as:

$$F_1 = 2 \frac{P \cdot R}{P + R} \quad (2)$$

where  $P$  is the precision and  $R$  is the recall, defined, for a binary classification, in (3) and (4) respectively:

$$P = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (3)$$

$$R = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (4)$$

For each condition, the precision is the relative fraction of correctly assigned instances amongst all the cases assigned to that specific class, while recall is the fraction of correctly recognised instances over the total amount of instances belonging to the same class (Apte et al. 1994). Such definitions are valid for binary classification, when the algorithm is challenged to identify the belonging of a certain dataset on the positive or negative classes (medical diagnosis is an example). In cases of multiclassification, the F1-score is calculated for each condition and successively weighted. Therefore, in this work the comparison is made on the average weighted F1-score (Al-Salemi et al. 2018): the closer F1-score is to 1 the more reliable and precise is its prediction power.

The process of training and testing the machine learning algorithm is repeated for ten times, where for each repetition, the datasets included in the training and testing are

randomly selected. This is done to avoid bias in evaluating the algorithm accuracy that may be achieved by selecting a “lucky” training dataset. The reported results are the average values obtained in the ten repetition steps.

In this study, two test case datasets were acquired, one at the same conditions of gas flow rate and impeller speed with the bi-phasic mixture, and a second one in which solids are added to the mixture.

## Results and discussion

### Choice of input parameters

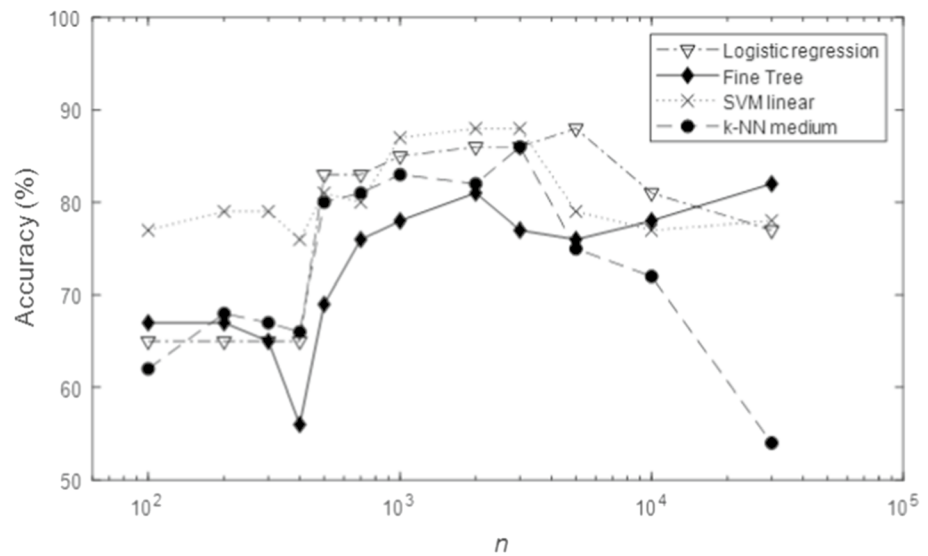
One of the potential issues for the application of AE as a real time tool is the computational time; it is essential that the analysis can be carried out in short time to evaluate the operating condition instantaneously. In this study, the parameters used by the algorithms to recognise the mixing regime are the spectrum points: the amplitude of each frequency present in the spectrum.

However, AE are usually acquired at high frequency resulting in a large amount of data; this can cause delay in processing and at the same time represents a challenge for the algorithms to avoid biasing and overfitting. For this reason, an initial study is carried out by manipulating  $n$ , the number of used features, in the spectrum for the classification. After ranking the frequency spectrum based on decreasing variance, a variable number,  $n$ , of frequencies is fed to the training process of the different algorithms. The obtained methods are then challenged to correctly classify the cross-validation dataset; as done for the training dataset, the corresponding  $n$  of the cross-validation spectra are used for this optimisation step. For this initial test, carried out on the biphasic case, one algorithm of each family is used for the analysis. The objective is to identify the optimal number of features that would represent a good compromise of obtained accuracy and processing time for the different algorithms. The results, in terms of accuracy, are shown in Fig. 6.

Figure 6 highlights how critical the choice of the parameter  $n$  is in ensuring good prediction performance. While the SVM starts already with performance around 75–80% even with a small number of features ( $n = 100$ ) with a local maximum in the range between 1000 and 3000 features, the other three algorithms perform significantly worse at low number of features. SVM Their accuracy progressively increase with increasing  $n$ , until the same range of SVM maximum is reached. In this condition all performance, except the decision tree, are in the same order of SVM. By increasing the number of features over 10,000, all the algorithms, except the decision tree, perform worse. This is due to the fact that when increasing the vector space dimension, the algorithms



**Fig. 6** Accuracy over the cross-validation dataset varying the number of  $n$  number of features for logistic regression, decision tree (fine), SVM (with linear kernel) and k-NN (with  $k=10$ ) in the gas–liquid case



**Table 2** Training time for fine tree, SVM (linear kernel), k-NN and the average value for the two-phase case

$n$	Training time [s]			
	Fine tree	SVM linear	k-NN medium	Average
100	0.41	0.18	0.089	0.2263
200	0.47	0.207	0.115	0.264
500	0.53687	0.336	0.1803	0.351
1000	0.77	0.544	0.361	0.558
2000	1.25	1.03	0.8016	1.02
5000	3.81	3.608	3.54	3.652
30,000	106.5	124.06	106.83	112.46

may be misguided by unimportant parts of the spectrum that tend to “homogenise” the acquired data at different conditions. The decision tree is less affected mainly because it is not based on geometrical parameters. The reason for SVM performing better than the others even at low number of features can be found in the data structure in the investigated vector space. The three classes can be differentiated by the SVM linear kernel with an accuracy around 80% at low number of features. Training time is reported in Table 2 for the different data points for the three methods built in Matlab® Classification learner (the logistic regression algorithm was not considered in this comparison because it was manually coded by the authors).

When  $n$  is increased over 1000, the processing time takes longer than 1 s; moreover,  $n = 1000$  corresponds to the maximum region of three of the four tested methods (Fig. 6), therefore, this is identified as the operating number of features for further analysis.

Although the optimisation is computed for the two-phase case, the training time for the three-phase case stays under 1 s as reported in Table 3.

**Table 3** Training time for fine tree, SVM (linear kernel), k-NN and the average value for the three-phase case

$n$	Training time [s]			
	Fine tree	SVM linear	k-NN medium	Average
1000	0.831	0.614	0.462	0.636

**Two-phase test**

The training was carried out for the chosen number of features based on the training dataset acquired for the biphasic condition. The test dataset, unseen by the machine in previous step was used to obtain the prediction performance of the set of chosen algorithms. Results in terms of accuracy for each algorithm at different regimes are reported in Table 4.

In terms of accuracy, all the methods exception made for the Coarse KNN have accuracy over 84%. Among the tested algorithms, the SVM with quadratic and cubic kernels seem to outperform the linear based SVM and the other algorithms, with accuracy equal to or higher than 90% in the three investigated conditions. Logistic regression also presents a high accuracy compared to other methods, while it is interesting to observe that k-NN performance inflects as the number of considered  $k$  neighbours is increased from 1 (fine  $k$ -NN) to 100 (coarse  $k$ -NN). In cases of classification problems, accuracy is often not enough to correctly describe the performance of the algorithm, therefore, for further comparison a weighted F1-score (Powers 2011) is calculated for the different algorithms as an average among the 10 repetitions. It is showed in Table 5.

Quadratic and cubic SVM achieve the highest result in F1-score, confirming what has already been observed from the accuracy comparison. By way of example, a parity plot from one of the repetition, showing the correct predictions

and the failures of the SVM cubic algorithm is reported in Fig. 7. The points on the identity line identify the correct assignments, while the misclassifications are in one of the other regions of the plot. Data show one of the repetitions (the median dataset) reported in Table 4.

All the ungasged cases are correctly identified by the algorithm, while three datasets for both loading and complete dispersion are misclassified. The occurrence of similar misclassifications can be due both to limitations of the model, but at the same time also to temporary oscillations of the mixing condition between the two regimes. Indeed, the three points belong to the points are at the edge between

the two regimes, respectively  $N=600$  rpm for loading and  $N=800$  rpm for complete dispersion (both at a gas feed rate of 1.5 vvm).

### Three-phase test

In the second part of the study, stainless-steel particles are added at different concentration and the acoustic signal is acquired at the three regimes. The same algorithms used for the biphasic system are applied for assessing the ML capability in this case, maintaining unchanged also

**Table 4** Average accuracy of the machine learning algorithms for the different regimes in the gas–liquid case

Regime	Method accuracy (%)								
	Logistic regr.	Fine tree	Medium tree	Linear SVM	Quad. SVM	Cubic SVM	Fine KNN	Medium KNN	Coarse KNN
Ungassed	89	93.1	93.1	86.7	94.9	96.1	100	84	88.4
Loading	85	87.8	87.8	83.2	97.1	97.1	89.2	83.2	30
Complete dispersion	94.7	84.0	84.0	97.6	87.6	90	84.0	82	0
Global	$90.2 \pm 2.9$	$88.9 \pm 4.2$	$88.9 \pm 4.2$	$86.7 \pm 3.8$	$93.8 \pm 3.8$	$94.4 \pm 3.7$	$91.6 \pm 3.9$	$83.8 \pm 5.10$	$44 \pm 8.6$

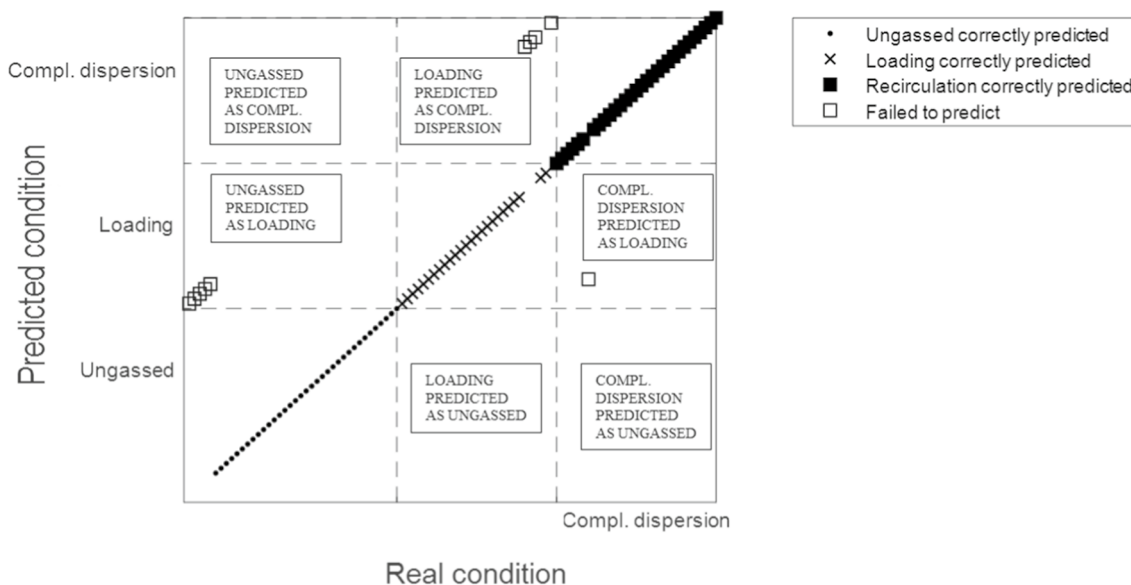
The last row reports the overall accuracy together with the standard deviation obtained in the 10 repetitions

**Table 5** Weighted F1 score obtained in the gas–liquid regime prediction case

Regime	Logistic regr.	Fine tree	Medium tree	Linear SVM	Quad. SVM	Cubic SVM	Fine KNN	Medium KNN	Coarse KNN*
F1 score	0.914	0.892	0.892	0.845	0.938	0.951	0.911	0.841	–

The value is averaged over the 10 repetitions

\*No value was calculated for the Coarse k-NN because of a null denominator



**Fig. 7** Parity plot reporting SVM quadratic results in classifying the different test datasets in the biphasic case

$n$ , the number of used features. The obtained results in terms of overall accuracy and F1-score are summarised in Table 6.

In the three-phase cases the obtained performance of all algorithms are high, with fine and medium k-NN achieving comparable performance to the cubic and quadratic SVM, (still being the highest performing algorithms) and the logistic regression algorithms. Figure 8 shows the parity plot obtained in the three-phase case. Data show one of the repetitions (the median dataset) reported in Table 6.

This result confirms that as the presence of solids, at low concentration, does not have major effects on the gas–liquid sparging dynamics within the tank, in similar fashion it does not significantly affect the ability of the ML tools to infer acoustic emission features to characterise the bi-phasic mixture. The algorithms are able to be trained to recognise the acoustic emission caused by the bubbles presence and interaction with the other phases and therefore it is possible for the system to identify the corresponding regime.

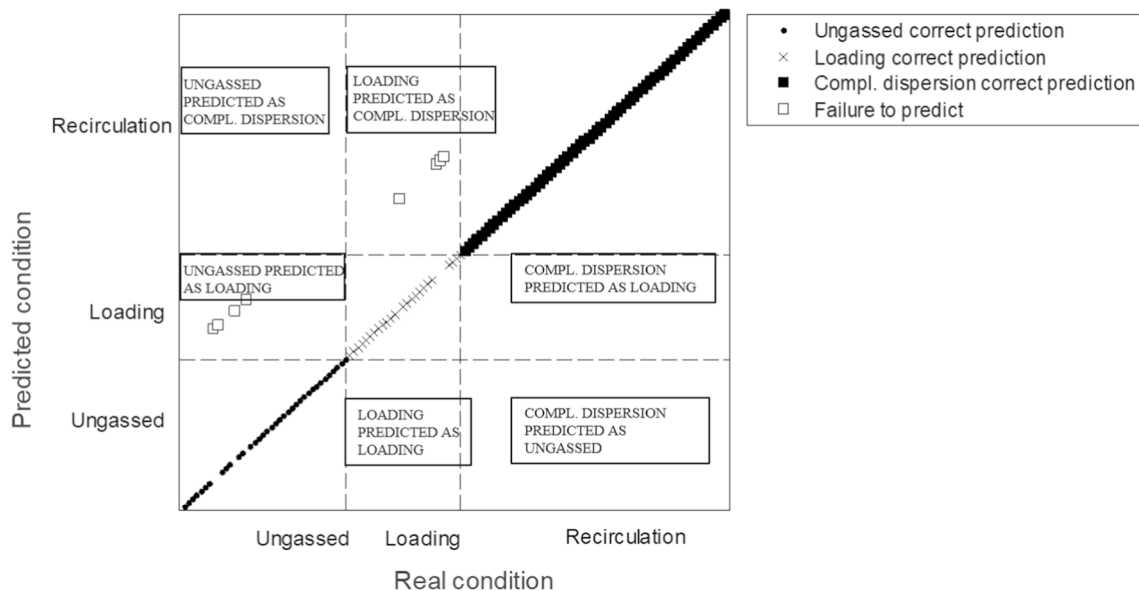
### Dynamic condition

To further challenge the technique, and in order to verify its flexibility over deviations from an initial well-defined condition, an additional test was run. The trained algorithm with the closest performance to the median in the gas–liquid condition is challenged to recognise the corresponding regimes in the three-phase condition. Such a situation is realistic in cases in which a system is characterised in a certain starting bi-phasic condition and then during the operation dynamic, a third phase is generated by reaction or addition to the mixture. Therefore, the acquired data used for the three-phase experiment are fed to the biphasic algorithm. The results, reported in Table 7, show the significant decrease in accuracy and F1-score exhibited by all the algorithms except the logistic regression. The influence of the solids in the acoustic spectrum indeed does not allow the system to recognise the correct gas–liquid regime. This represents a limit to the flexibility of the geometrical based algorithms (SVM and k-NN) and the decision tree that clearly do not succeed in correlating characteristic features in the biphasic

**Table 6** Overall accuracy and weighted F1-score for the different algorithms in gas–solid–liquid regime prediction

Regime	Logistic regr.	Fine tree	Medium tree	Linear SVM	Quad. SVM	Cubic SVM	Fine KNN	Medium KNN	Coarse KNN
Overall accuracy (%)	90.1 ± 3.4	83.1 ± 4.2	83.8 ± 4.9	78.8 ± 4.6	93.8 ± 4.2	94.2 ± 3.9	92.5 ± 2.9	90.9 ± 3.4	56.5 ± 9.5
Weighted F1-score	0.931	0.852	0.881	0.813	0.939	0.941	0.912	0.888	0.734

Data are calculated over 10 repetitions



**Fig. 8** Parity plot reporting SVM quadratic results in classifying the different signals in the gas–solid–liquid test

**Table 7** Overall accuracy and weighted F1-score for the different algorithms in gas–solid–liquid regime prediction using the algorithms trained in the gas–liquid condition

Algorithm	Logistic regr.	Fine Tree	Medium Tree	Linear SVM	Quad. SVM	Cubic SVM	Fine KNN	Medium KNN	Coarse KNN
Overall accuracy (%)	90.8	32.4	42.3	33.0	34.4	27.3	49.1	51.2	29.8
Weighted F1-score	0.928	0.496	0.539	0.431	0.459	0.438	0.631	0.583	0.390

case with corresponding ones in the three-phase case. The logistic regression instead is able to present performances similar to the ones obtained in the previous two cases. The reason is related to the different nature of this algorithm: in each case a probability of belonging of the instance to the three different cases is calculated independently. Although all these probabilities decrease, the correct one outranks the others and therefore the system is able to correctly predict the gas–liquid condition.

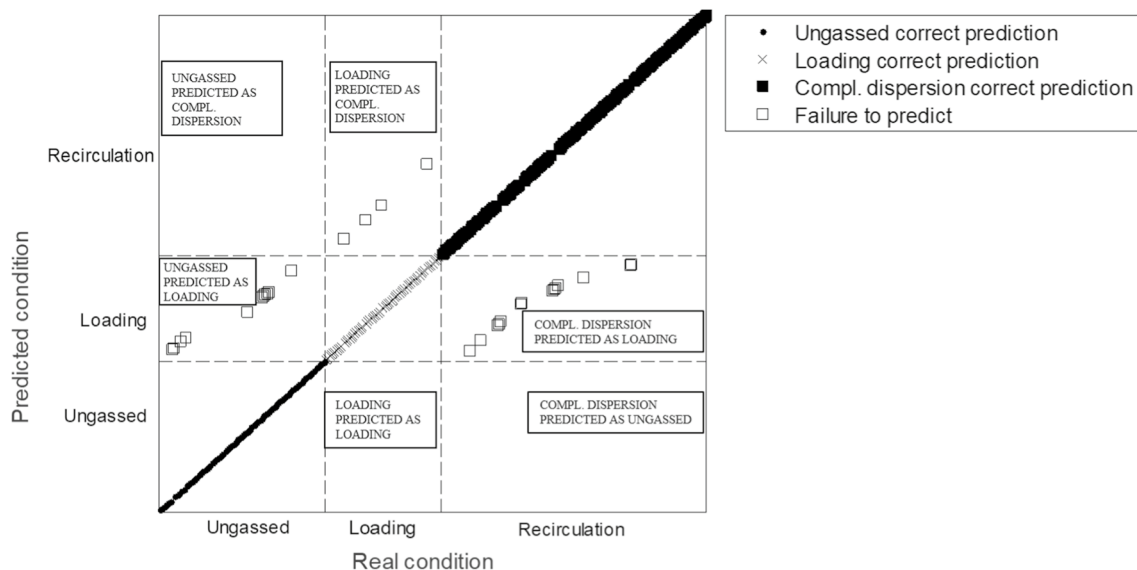
Figure 9 reports the parity plot for the best case (logistic regression), and does not show any particular clustered dataset presenting fallacious behaviour. Amongst the processed data 91.2% are correctly recognised by the ML.

The logistic regression, already well-performing in the other cases demonstrates higher flexibility than the other machine learning algorithms investigated in this study. This can be certainly related to the system used and the choice of data-processing algorithm remains an important issue to be addressed. Deviations from ideal conditions is an everyday challenge on process plants and the ability to adapt to non-forecasted conditions is a critical characteristic for an implementable measurement technique.

## Conclusions

In this work, a methodology to use Acoustic Emission data in combination with machine learning algorithms, to identify gas–liquid mixing conditions in a stirred tank was proposed. The machine was trained to recognise ungasged, loading and recirculation condition using a training dataset, consisting in the acoustic spectrum acquired at the different regimes in two-phase (gas–liquid) and three-phase (gas–solid–liquid) stirred tank. The training procedure was repeated ten times for each algorithm by varying the data points in the leaning dataset. Average results show that with appropriate training, some of the investigated algorithms (SVM with quadratic and cubic kernel and logistic regression) were able to correctly recognise the regimes corresponding to the tested spectra with accuracy higher than 90%.

The system was then challenged to make prediction on the gas sparging regime in the three-phase runs while using the learning gained in the bi-phasic condition to evaluate flexibility of the method in dynamic conditions

**Fig. 9** Parity plot reporting logistic regression results in classifying the different signals in the gas–solid–liquid test using the gas–liquid trained algorithms

(as for example in precipitation batch processes). Most of the algorithms did not successfully address this task, presenting accuracy in some cases significantly lower than 50%; however logistic regression, above all, was the one that is able to perform at the same level as for previous tests (> 90%).

The conducted study aimed to investigate AE as a potential diagnostic and condition monitoring technique for fluid mixing applications in combination with the use of machine learning algorithms and proposes a methodology to evaluate the most appropriate algorithms. Although amongst the wide range of algorithms available in the literature, only some were taken into consideration, from this investigation it is possible to infer that a combined use of AE and ML could represent a powerful tool for in situ monitoring of two-phase and three-phase mixing.

**Acknowledgements** Giuseppe Forte is an EngD student at the University of Birmingham funded jointly by the Engineering & Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Formulation Engineering (EP/L015153/1) and Johnson Matthey.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Addali, A., Al-lababidi, S., Yeung, H., Mba, D., & Khan, F. (2010). Acoustic emission and gas-phase measurements in two-phase flow. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 224(4), 281–290. <https://doi.org/10.1243/09544089JPME359>.
- Aldrich, C., & Theron, D. A. (2000). Acoustic estimation of the particle size distributions of sulphide ores in a laboratory ball mill. *Journal of the South African Institute of Mining and Metallurgy*, 100(4), 243–248.
- Al-Salemi, B., Ayob, M., & Noah, S. A. M. (2018). Feature ranking for enhancing boosting-based multi-label text categorization. *Expert Systems with Applications*, 113, 531–543. <https://doi.org/10.1016/j.eswa.2018.07.024>.
- Apte, C., Damerau, F., & Weiss, S. (1994). Automated learning of decision rules for text categorization. *ACM Transactions on Information Systems*, 12(3), 233–251. <https://doi.org/10.1145/183422.183423>.
- Bañares-Alcantara, R., Venkatasubramanian, V., Westerberg, A. W., & Rychener, M. D. (1985). Knowledge-based expert systems: An emerging technology for CAD in chemical engineering. *Chemical Engineering Progress*, 9, 127–142.
- Bañares-Alcantara, R., Westerberg, A. W., Ko, E. I., & Rychener, M. D. (1987). Decade—A hybrid expert system for catalyst selection—I. Expert system consideration. *Computers & Chemical Engineering*, 11(3), 265–277. [https://doi.org/10.1016/0098-1354\(87\)85008-1](https://doi.org/10.1016/0098-1354(87)85008-1).
- Bao, Y. Y., Hao, Z. G., Gao, Z. M., Shi, L. T., Smith, J. M., & Thorpe, R. B. (2006). Gas dispersion and solid suspension in a three-phase stirred tank with multiple impellers. *Chemical Engineering Communications*, 193(7), 801–825. <https://doi.org/10.1080/00986440500267261>.
- Bender, R., & Kuss, O. (2010). Methods to calculate relative risks, risk differences, and numbers needed to treat from logistic regression. *Journal of Clinical Epidemiology*, 63(1), 7–8. <https://doi.org/10.1016/j.jclinepi.2009.07.007>.
- Boyd, J. W. R., & Varley, J. (2001). The uses of passive measurement of acoustic emissions from chemical engineering processes. *Chemical Engineering Science*, 56(5), 1749–1767. [https://doi.org/10.1016/S0009-2509\(00\)00540-6](https://doi.org/10.1016/S0009-2509(00)00540-6).
- Bragg Sir, W. H. (1921). *The world of sound*. London: Bell.
- Caesarendra, W., Widodo, A., & Yang, B.-S. (2010). Application of relevance vector machine and logistic regression for machine degradation assessment. *Mechanical Systems and Signal Processing*, 24, 1161–1171. <https://doi.org/10.1016/j.ymssp.2009.10.011>.
- Cents, A. H. G., Brilmant, D. W. F., & Versteeg, G. F. (2005). Ultrasonic investigation of hydrodynamics and mass transfer in a gas–liquid(–liquid) stirred vessel. *International Journal of Chemical Reactor Engineering*. <https://doi.org/10.2202/1542-6580.1164>.
- Chapman, C. M., Nienow, A. W., Cooke, M., & Middleton, J. C. (1983). Particle–gas–liquid mixing in stirred vessels. Part IV: Mass transfer and final conclusions. *Chemical Engineering Research and Design*, 61(3), 182–185.
- Chen, B. J., Chen, X. F., & Li, B. (2011). Reliability estimation for cutting tool based on logistic regression model. *Chinese Journal of Mechanical Engineering*, 47(18), 158–164.
- Cole, P., & Watson, J. (2005). Acoustic emission for corrosion detection. *Presented at the Middle East Nondestructive Conference & Exhibition, Bahrain, Manama* (p. 7).
- Cooley, J. W., Lewis, P. A. W., & Welch, P. D. (1969). The fast Fourier transform and its applications. *IEEE Transactions on Education*, 12(1), 27–34. <https://doi.org/10.1109/TE.1969.4320436>.
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>.
- Crouter, A., & Briens, L. (2015). Passive acoustic emissions from particulates in a V-blender. *Drug Development and Industrial Pharmacy*, 41(11), 1809–1818. <https://doi.org/10.3109/03639045.2015.1009913>.
- Elforjani, M., & Shanbr, S. (2018). Prognosis of bearing acoustic emission signals using supervised machine learning. *IEEE Transactions on Industrial Electronics*, 65(7), 5864–5871. <https://doi.org/10.1109/TIE.2017.2767551>.
- Esbensen, K. H., Halstensen, M., Tønnesen Lied, T., Saudland, A., Svalestuen, J., de Silva, S., et al. (1998). Acoustic chemometrics—From noise to information. *Chemometrics and Intelligent Laboratory Systems*, 44(1–2), 61–76. [https://doi.org/10.1016/S0169-7439\(98\)00114-2](https://doi.org/10.1016/S0169-7439(98)00114-2).
- Ferizi, U., Besser, H., Hysi, P., Jacobs, J., Rajapakse, C. S., Chen, C., et al. (2019). Artificial intelligence applied to osteoporosis: A performance comparison of machine learning algorithms in predicting fragility fractures from MRI data. *Journal of Magnetic Resonance Imaging*, 49(4), 1029–1038. <https://doi.org/10.1002/jmri.26280>.
- Ford, J. J., Heindel, T. J., Jensen, T. C., & Drake, J. B. (2008). X-ray computed tomography of a gas-sparged stirred-tank reactor. *Chemical Engineering Science*, 63(8), 2075–2085. <https://doi.org/10.1016/j.ces.2008.01.007>.



- Forte, G., Alberini, F., Simmons, M. J. H., & Stitt, E. H. (2019). Measuring gas hold-up in gas–liquid/gas–solid–liquid stirred tanks with an electrical resistance tomography linear probe. *AIChE Journal*, 65(6), e16586. <https://doi.org/10.1002/aic.16586>.
- Gao, Z., Smith, J. M., & Müller-Steinhagen, H. (2001). Gas dispersion in sparged and boiling reactors. *Chemical Engineering Research and Design*, 79(8), 973–978. <https://doi.org/10.1205/02638760152721523>.
- Griffin, J., & Chen, X. (2016). Real-time simulation of neural network classifications from characteristics emitted by acoustic emission during horizontal single grit scratch tests. *Journal of Intelligent Manufacturing*, 27, 507–523. <https://doi.org/10.1007/s10845-014-0883-x>.
- Hilbe, J. M. (2009). *Logistic regression models* (1st ed., Vol. 1). Chapman and Hall CRC Press. Retrieved April 11, 2019, from <https://www.crcpress.com/Logistic-Regression-Models/Hilbe/p/book/9781138106710>.
- Hu, H., Wang, L., & Luh, P. (2015). Intelligent manufacturing: New advances and challenges. *Journal of Intelligent Manufacturing*, 26, 841–843. <https://doi.org/10.1007/s10845-015-1148-z>.
- Jamshed, A., Cooke, M., Ren, Z., & Rodgers, T. L. (2018). Gas–liquid mixing in dual agitated vessels in the heterogeneous regime. *Chemical Engineering Research and Design*, 133, 55–69. <https://doi.org/10.1016/j.cherd.2018.02.034>.
- Jang, J.-R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665–685. <https://doi.org/10.1109/21.256541>.
- Karamzadeh, S., Abdullah, S. M., Halimi, M., Shayan, J., & Rajabi, M. J. (2014). Advantage and drawback of support vector machine functionality. In *2014 international conference on computer, communications, and control technology (I4CT)* (pp. 63–65). <https://doi.org/10.1109/i4ct.2014.6914146>.
- Kasemsumran, P., Auephanwiriyaikul, S., & Theera-Umpon, N. (2016). Face recognition using string grammar fuzzy K-nearest neighbor. In *2016 8th international conference on knowledge and smart technology (KST)* (pp. 55–59). <https://doi.org/10.1109/kst.2016.7440531>.
- Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31(3), 249–268.
- Kouroussis, D., Anastassopoulos, A., Vionis, P., & Kolovos, V. (2000). Unsupervised pattern recognition of acoustic emission from full scale testing of a wind turbine blade. *Journal of Acoustic Emission*, 18, 217–223.
- Li, C., Sanchez, R.-V., Zurita, G., Cerrada, M., Cabrera, D., & Vásquez, R. E. (2016). Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mechanical Systems and Signal Processing*, 76–77, 283–293. <https://doi.org/10.1016/j.ymssp.2016.02.007>.
- Li, H., Wang, Y., Zhao, P., Zhang, X., & Zhou, P. (2015). Cutting tool operational reliability prediction based on acoustic emission and logistic regression model. *Journal of Intelligent Manufacturing*, 26(5), 923–931. <https://doi.org/10.1007/s10845-014-0941-4>.
- Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance*, 1(3), 249–276. [https://doi.org/10.1016/0378-4266\(77\)90022-X](https://doi.org/10.1016/0378-4266(77)90022-X).
- Middleton, J. C. (1992). Gas–liquid dispersion and mixing. In N. Harnby, M. F. Edwards, & A. W. Nienow (Eds.), *Mixing in the process industries* (pp. 322–363). Oxford: Butterworth-Heinemann. <https://doi.org/10.1016/b978-075063760-2/50036-6>.
- Murthy, S. K. (1998). Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Mining and Knowledge Discovery*, 2(4), 345–389. <https://doi.org/10.1023/A:1009744630224>.
- National Academies of Sciences, E. (2018). *Data science: Opportunities to transform chemical sciences and engineering: Proceedings of a Workshop—In brief*. <https://doi.org/10.17226/25191>.
- Nienow, A. W., Warmoeskerken, M. M. C. G., Smith, J. M., & Konno, M. (1985). On the flooding/loading transition and the complete dispersal condition in aerated vessels agitated by a Rushton-turbine. In *5th European conference on mixing* (pp. 143–154). Wurtsburg, West Germany.
- Nordon, A., Carella, Y., Gachagan, A., Littlejohn, D., & Hayward, G. (2006). Factors affecting broadband acoustic emission measurements of a heterogeneous reaction. *The Analyst*, 131(2), 323–330. <https://doi.org/10.1039/B510922A>.
- Nordon, A., Waddell, R. J. H., Bellamy, L. J., Gachagan, A., McNab, D., Littlejohn, D., et al. (2004). Monitoring of a heterogeneous reaction by acoustic emission. *The Analyst*, 129(5), 463–467. <https://doi.org/10.1039/b402875a>.
- Pomponi, E., & Vinogradov, A. (2013). A real-time approach for acoustic emission clustering. *Mechanical Systems and Signal Processing*, 40(2), 791–804. <https://doi.org/10.1016/j.ymssp.2013.03.017>.
- Powers, D. M. W. (2011). Evaluation: From precision, recall and fmeasure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Prasath, V. B. S., Alfeilat, H. A. A., Lasassmeh, O., & Hassanat, A. B. A. (2017). *Distance and similarity measures effect on the performance of K-nearest neighbor classifier—A review*. [arXiv:1708.04321](https://arxiv.org/abs/1708.04321) [cs].
- Rink, B., Harabagiu, S., & Roberts, K. (2011). Automatic extraction of relations between medical concepts in clinical texts. *Journal of the American Medical Informatics Association*, 18(5), 594–600. <https://doi.org/10.1136/amiajnl-2011-000153>.
- Rizwan, M., & Anderson, D. V. (2014). Using k-nearest neighbor and speaker ranking for phoneme prediction. In *2014 13th international conference on machine learning and applications* (pp. 383–387). <https://doi.org/10.1109/icmla.2014.68>.
- Rozak, P., Zielinski, J., Czop, P., Jablonski, A., Barszcz, T., & Mareczek, M. (2018). Supervised classification methods in condition monitoring of rolling element bearings. In A. Timofiejczuk, F. Chaari, R. Zimroz, W. Bartelmus, & M. Haddar (Eds.), *Advances in condition monitoring of machinery in non-stationary operations, Cmmno 2016* (Vol. 9). Berlin: Springer.
- Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3), 660–674. <https://doi.org/10.1109/21.97458>.
- Soofi, A. A., & Awan, A. (2017). Classification techniques in machine learning: Applications and issues. *Journal of Basic and Applied Sciences*, 13, 459–465.
- Spellings, M., & Glotzer, S. C. (2018). Machine learning for crystal identification and discovery. *AIChE Journal*, 64(6), 2198–2206. <https://doi.org/10.1002/aic.16157>.
- Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9(3), 293–300. <https://doi.org/10.1023/A:1018628609742>.
- Ulissi, Z. W., Tang, M. T., Xiao, J., Liu, X., Torelli, D. A., Karamad, M., et al. (2017). Machine-learning methods enable exhaustive searches for active bimetallic facets and reveal active site motifs for CO<sub>2</sub> reduction. *ACS Catalysis*, 7(10), 6600–6608. <https://doi.org/10.1021/acscatal.7b01648>.
- Vapnik, V. (2000). *The nature of statistical learning theory* (2nd ed.). New York: Springer.
- Veera, P. U., Patwardhan, A. W., & Joshi, J. B. (2001). Measurement of gas hold-up profiles in stirred tank reactors by gamma ray attenuation technique. *Chemical Engineering Research and Design*, 79(6), 684–688. <https://doi.org/10.1205/026387601316971352>.



- Venkatasubramanian, V. (2019). The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal*, *65*(2), 466–478. <https://doi.org/10.1002/aic.16489>.
- Warmoeskerken, M. M. C. G., & Smith, J. M. (1985). Flooding of disc turbines in gas–liquid dispersions: A new description of the phenomenon. *Chemical Engineering Science*, *40*(11), 2063–2071. [https://doi.org/10.1016/0009-2509\(85\)87023-8](https://doi.org/10.1016/0009-2509(85)87023-8).
- Whitaker, M., Baker, G. R., Westrup, J., Goulding, P. A., Rudd, D. R., Belchamber, R. M., et al. (2000). Application of acoustic emission to the monitoring and end point determination of a high shear granulation process. *International Journal of Pharmaceutics*, *205*(1), 79–91. [https://doi.org/10.1016/S0378-5173\(00\)00479-8](https://doi.org/10.1016/S0378-5173(00)00479-8).
- Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. <https://doi.org/10.1007/s10115-007-0114-2>.
- Wu, W.-Z., & Zhang, W.-X. (2002). Neighborhood operator systems and approximations. *Information Sciences*, *144*(1), 201–217. [https://doi.org/10.1016/S0020-0255\(02\)00180-9](https://doi.org/10.1016/S0020-0255(02)00180-9).
- Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23–45. <https://doi.org/10.1080/21693277.2016.1192517>.
- Xue, D., Balachandran, P. V., Hogden, J., Theiler, J., Xue, D., & Lookman, T. (2016). Accelerated search for materials with targeted properties by adaptive design. *Nature Communications*, *7*, 11241. <https://doi.org/10.1038/ncomms11241>.
- Yan, J., Koç, M., & Lee, J. (2004). A prognostic algorithm for machine performance assessment and its application. *Production Planning & Control*, *15*(8), 796–801. <https://doi.org/10.1080/09537280412331309208>.
- Yan, J., & Lee, J. (2004). Degradation assessment and fault modes classification using logistic regression. *Journal of Manufacturing Science and Engineering*, *127*(4), 912–914. <https://doi.org/10.1115/1.1962019>.
- Yawalkar, Archis A., Pangarkar, Vishwas G., & Beenackers Anthony, A. C. M. (2002). Gas hold-up in stirred tank reactors. *The Canadian Journal of Chemical Engineering*, *80*, 158–166. <https://doi.org/10.1002/cjce.5450800117>.
- Zhang, L., Liu, Q., Yang, W., Wei, N., & Dong, D. (2013). An improved K-nearest neighbor model for short-term traffic flow prediction. *Procedia—Social and Behavioral Sciences*, *96*, 653–662. <https://doi.org/10.1016/j.sbspro.2013.08.076>.
- Zhao, W., Buffo, A., Alopaeus, V., Han, B., & Louhi-Kultanen, M. (2017). Application of the compartmental model to the gas–liquid precipitation of CO<sub>2</sub>-Ca(OH)<sub>2</sub> aqueous system in a stirred tank. *AIChE Journal*, *63*(1), 378–386. <https://doi.org/10.1002/aic.15567>.
- Zhao, W., Han, B., Jakobsson, K., Louhi-Kultanen, M., & Alopaeus, V. (2016). Mathematical model of precipitation of magnesium carbonate with carbon dioxide from the magnesium hydroxide slurry. *Computers & Chemical Engineering*, *87*, 180–189. <https://doi.org/10.1016/j.compchemeng.2016.01.013>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.