A MACHINE LEARNING APPROACH TO CLASSIFICATION OF GAS ENTRAINMENT AND IMPELLER WEAR IN CENTRIFUGAL PUMPS

by

BRYAN BOHN
B.S., University of Connecticut, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND DOCTORAL STUDIES

(Mechanical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February, 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

A MACHINE LEARNING APPROACH TO CLASSIFICATION OF GAS ENTRAINMENT AND IMPELLER WEAR IN CENTRIFUGAL PUMPS

submitted by Bryan Bohn in partial fulfillment of the requirements for

the degree of Master of Applied Science

in Mechanical Engineering

Examining Committee:

Boris Stoeber, Professor, Mechanical Engineering and Electrical and Computer Engineering, UBC
Supervisor

Bhushan Gopaluni, Professor, Chemical and Biological Engineering, UBC
Supervisory Committee Member

Hongshen Ma, Professor, Mechanical Engineering, UBC
Supervisory Committee Member

Minkyun Noh, Assistant Professor, Mechanical Engineering, UBC
Supervisory Committee Member

ABSTRACT

Centrifugal pumps are a fundamental part of fluid transport around the world. Consequently, they are also one of the world's dominant energy consumers. The impacts of inefficient operation and undiagnosed wear are widely documented and can be disastrous environmentally, financially, and logistically. Though commercial tools and methods for monitoring pump performance are abundant, they are used infrequently in practice. This phenomenon derives from several factors, including monitoring systems' poor scalability to function with large numbers of pumps, acquisition costs, the necessity for additional technical personnel, and stringent policies constraining process downtime.

This thesis describes the development of an affordable, adaptable sensing method for classifying two conditions detrimental to centrifugal pump operation; gas entrainment and radial impeller wear. The method utilizes dynamic pressure measurements, collected at the pump discharge using a solitary, conventional pressure transducer. Decomposing these pressure fluctuations into a novel array of statistical features yields characteristic trends correlated to the target phenomena. These features are then used to train a series of machine learning algorithms, including multilayer perceptrons (MLP), support vector machines (SVM), and random forests, which are in turn used to characterize the target conditions using binary, multi-class, and regression methods.

Dynamic pressure data for training and testing the classification algorithms is generated using simulated and experimental methods. The binary MLP model predicts gas entrainment exceeding a 2% void fraction of air with 90% accuracy, and radial wear exceeding 1.5% of the impeller diameter with 97% accuracy. The multi-class MLP classifies gas entrainment and radial impeller wear into severity classes spanning 1% increments with 62% and 82% success rates, respectively. The random forest regression model achieves a median prediction error of 0.44% for gas entrainment and 0.16% for impeller wear.

The diagnostic system presented in this research is unique in that it is not conceived as a standalone tool for pump users, but rather a shared process to be trained and configured by the pump manufacturer, then implemented by the operators. In its envisioned application, the scope of the classified phenomena would be augmented by the manufacturer to capture a wide variety of pump performance characteristics.

LAY SUMMARY

Centrifugal pumps are one of the world's most important machines. They convert electrical power into fluid movement to transport liquids. In doing so, they also consume a lot of energy. When pumps wear down, or the fluid they are transporting behaves unexpectedly, they may start operating poorly, which wastes energy. There are a variety of commercial devices available to measure whether or not a centrifugal pump is operating well, but, in practice, they are not widely used because of their cost and difficulty to implement.

We developed a more practical method to quantify two kinds of problems that would cause a centrifugal pump to waste energy: mechanical wear and air bubbles in the fluid flow. Our technique uses a single sensor to measure pressure fluctuations in the fluid. After measuring fluctuations in many different conditions, we created machine learning algorithms to identify how severe the problem is.

PREFACE

This thesis is an original work by the author, Bryan Bohn, at The University of British Columbia, Vancouver, Canada (UBC). The research herein was performed in its entirety by the author, under the supervision and guidance of Boris Stoeber, Professor of Mechanical Engineering and Electrical and Computer Engineering, and Bhushan Gopaluni, Professor of Chemical and Biological Engineering.

The work presented in this thesis was conducted as part of the Energy Reduction in Mechanical Pulping (ERMP) program at UBC. The Principal Investigator for the ERMP is James Olson, Professor and Dean of the Faculty of Applied Science. Dean Olson conceived the original concept for the project and provided guidance and advising throughout.

A portion of this research was presented at the 2019 IEEE Sensors conference in Montreal, QC, Canada and published in the proceedings: *B. Bohn, J. Olson, B. Gopaluni, and B. Stoeber, "Sensing Concept for Practical Performance-Monitoring of Centrifugal Pumps," in 2019 IEEE SENSORS, Montreal, QC, Canada, Oct. 2019, pp. 1–4, doi: 10.1109/SENSORS43011.2019.8956559.* The publication discusses the motivation for this work and proposes a concept for pump performance monitoring using a multi-sensor approach. The remainder of the work in this thesis is unpublished, as of its date of submission.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol	Description	Units
C_{po}	Specific heat at constant pressure of the pump outlet	$J/(kg \cdot K)$
$C_{\varepsilon 1}$	Turbulent production rate scalar	Dimensionless
$C_{\varepsilon 2}$	Turbulent dissipation rate scalar	Dimensionless
C_{μ}	Turbulent eddy viscosity scalar	Dimensionless
d	Pipe diameter	m
D_o	Diameter of the unworn impeller	m
D_α	Diameter of the actual impeller	m
ΔD	Wear magnitude of the impeller	m
e	Euler's number `	Dimensionless
E_{FFT}	Energy of the frequency spectrum	Dimensionless
E_{R}	Energy of the autocorrelation signal	Dimensionless
E_{s}	Signal energy	Dimensionless
F	Volume shear force	N/m^3
$\boldsymbol{F}$	Feature vector	Dimensionless
$\{\boldsymbol{F}\}$	Feature space	Dimensionless
f_{s}	Sampling frequency	Hz
f_{max}	Maximum frequency of interest	Hz
f_{min}	Minimum frequency of interest	Hz
f_{nyq}	Nyquist frequency	Hz
Δf	Frequency resolution	Hz
g	Acceleration of gravity	m/s^2
$\boldsymbol{g}$	Gravitational field	m/s^2
G_{s}	Signal skewness	Dimensionless
H	Head	m
$\boldsymbol{I}$	Identity tensor	Dimensionless
k	Turbulent kinetic energy	$J/kg, m^2/s^2$
K_{s}	Signal kurtosis	Dimensionless
L	Length	m
$\dot{m}$	Mass flow rate	kg/s
M_{air}	Number of prediction classes for air entrainment	Dimensionless
M_{wear}	Number of prediction classes for air impeller wear	Dimensionless
N	Number of sampled data points	Dimensionless
N_{imp}	Number of impeller blades	Dimensionless
p	Fluid pressure	Pa
$\boldsymbol{P}$	Prediction vector	Dimensionless
p_{in}	Fluid pressure at the pump inlet	Pa
P_k	Turbulent kinetic energy production	$J/(m^3 \cdot s)$
p_{out}	Fluid pressure at the pump outlet	Pa
$\boldsymbol{P}_{\mathrm{out}}$	Signal vector of discharge pressure measurements	kPa
$\widehat{\boldsymbol{P}}_{\mathrm{out}}$	Unbiased signal vector of discharge pressure measurements	kPa
$\widetilde{\boldsymbol{P}}_{\mathrm{out}}$	Normalized vector of discharge pressure measurements	Dimensionless
Δp	Average pressure across the pump	Pa
Δp_{L}	Pressure loss over a pipe of length L	Pa

Q	Volumetric flow rate	m^3/s
R_l	Autocorrelation coefficient	Dimensionless
R^2	Coefficient of determination	Dimensionless
S	Strain rate tensor	$J/(m^3 \cdot s)$
t	Time	s
T_{in}	Temperature at the pump inlet	K
T_{out}	Temperature at the pump outlet	K
t_{s}	Sampling interval	s
T_{s}	Sample time duration	s
ΔT	Temperature change across the pump	K
ΔT_{i}	Temperature change across an ideal, isentropic pump	K
$\boldsymbol{u}$	Flow field velocity vector	m/s
v	Specific volume	m^3/kg
v_{in}	Flow velocity at the pump inlet	m/s
v_{out}	Flow velocity at the pump outlet	m/s
V_{s}	Signal variance	Dimensionless
w	Neural network weight	Dimensionless
W_{fluid}	Fluid power output from the pump	W
W_{i}	Input power required to operate an ideal, isentropic pump	W
W_{in}	Input power to the pump	W
X	Frequency magnitude	Dimensionless
$\boldsymbol{X}$	FFT spectrum	Dimensionless
x_{n}	Pressure magnitude at index n	kPa
$\hat{x}_{\text{n}}$	Unbiased pressure magnitude at index n	kPa
$\tilde{x}_{\text{n}}$	Normalized pressure magnitude at index n	kPa
$\bar{\boldsymbol{y}}$	Weighted average vector	Dimensionless

α_{imp}	Impeller material loss ratio	Dimensionless
γ_{air}	Volumetric flow rate of air of air into the pump	m^3/s
γ_{w}	Volumetric flow rate of air of water into the pump	m^3/s
Γ	Prediction target value	Dimensionless
ε	Turbulent kinetic energy dissipation rate	$J/(kg \cdot s)$
ϵ	Level-set transition thickness parameter	Dimensionless
ζ	Level-set tuning parameter	Dimensionless
η_{conv}	Thermal efficiency determined using the conventional method	Dimensionless
η_{LR}	Learning rate coefficient	Dimensionless
η_{therm}	Thermal efficiency determined using the thermodynamic method	Dimensionless
θ	Angular position of the impeller	deg
μ	Dynamic viscosity	$(N \cdot s)/m^2$
μ_{T}	Turbulent eddy viscosity	$(N \cdot s)/m^2$
ν_{l}	Kinematic viscosity of the liquid phase	m^2/s
ρ	Average fluid density	kg/m^3
ρ_{air}	Average fluid density of air	kg/m^3
ρ_{water}	Average fluid density of liquid water	kg/m^3
σ_k	Adjustable constant	Dimensionless
σ_{ε}	Adjustable constant	Dimensionless

τ_{air}	Air entrainment classification threshold	Dimensionless
τ_{wear}	Impeller wear classification threshold	Dimensionless
ϕ	Level-set variable	Dimensionless
φ_{air}	Volume fraction of air	Dimensionless
ω	Rotating frequency of the impeller	Hz
$\omega_{\max}$	Maximum rotating frequency	Hz
$\omega_{\min}$	Minimum rotating frequency	Hz

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professor Boris Stoeber and Professor Bhushan Gopaluni for allowing me the opportunity to undertake this research. Your guidance, enthusiasm, good humor, and patience have made for an exceptional research experience. My time working with both of you has been engaging and illuminating, and for that I am tremendously appreciative. Thank you both for everything.

I also extend my gratitude to the additional members of my examining committee, Professor Hongshen Ma and Professor Minkyun Noh. Thank you for contributing your time and constructive feedback. I appreciate it immensely.

Additionally, I would like the thank Professor James Olson, Dean of the UBC Faculty of Applied Science and the originator of this research project, for contributing guidance, perspective, and positivity throughout. It has been a pleasure working with you and the ERMP members. Thank you for all your help.

My time at UBC has been outstanding. That is no small praise, considering that a year of it has overlapped a devastating global pandemic. I am thankful to the wonderful people in my lab for making the endeavor so gratifying, both academically and personally. I would like to acknowledge Sam, Pranav, Crystal, Nick, Sajana, Martina, Robin, Hamed, Jorge, and all the Stoeber Lab members, past and present. You have helped make a truly extraordinary and memorable experience for me. I'm looking forward to seeing you all in person again soon.

There have been many minor heroes whose efforts have benefited my work and life at UBC (sometimes without their knowing). Lest they go unrecognized, I would like to thank George Soong and Reanna Seifert at the Paper and Pulp Centre for their perpetual help resolving problems, the maintenance staff of the Fred Kaiser Building, the UBC Recreation employees at Thunderbird Arena, the serving staffs at the Gallery, JamJar, and Rain or Shine, the owners of Dentry's, my friends in BC and beyond, and my teammates on the Rain City Pigeons hockey team.

Lastly, I would like to acknowledge the people at home who sent their love and support across the whole continent while I uprooted to BC; Jer, Nicole, Robin, Corey, Sean, Conor, and the dogs – thank you guys. Finally, thank you mom and dad. I love you all.

To my most enthusiastic collaborator, Captain.

1. INTRODUCTION

1.1 Motivation

Centrifugal pumps have been a cornerstone of fluid-moving processes for centuries [1]. They are essential to virtually all modern industries, including manufacturing, mining, civil infrastructure, textiles, paper and pulp production, petrochemical refining, construction, and power production. At an applied level, they perform myriad functions, like moving cooling fluid through web servers, extracting drinking water from wells, or exchanging the contents of fermentation tanks at a brewery.

Whether a low wattage pump draining soapy water from a washing machine or a multi-thousand-horsepower colossus processing slurry from a mine, the mechanics of centrifugal pump operation are fundamentally similar. Kinetic energy is provided by a rotating shaft, which is typically driven by an electric motor or combustion engine. The rotating frequency (RF) of the shaft is designated ω. The shaft connects to a rotor, called an impeller, that revolves within the pump housing, imparting hydrodynamic energy on the incoming working fluid in the form of increased kinetic energy, pressure, and temperature. The impeller is comprised of one or more vanes, called "blades," which may be straight, curved, or a more complex helical shape in three dimensions. The total number of blades on a particular impeller is N_{imp}. Figure 1.1 shows the main components of a centrifugal pump system. The overall size, shape, construction materials, and configuration of a centrifugal pump depend on its intended application.

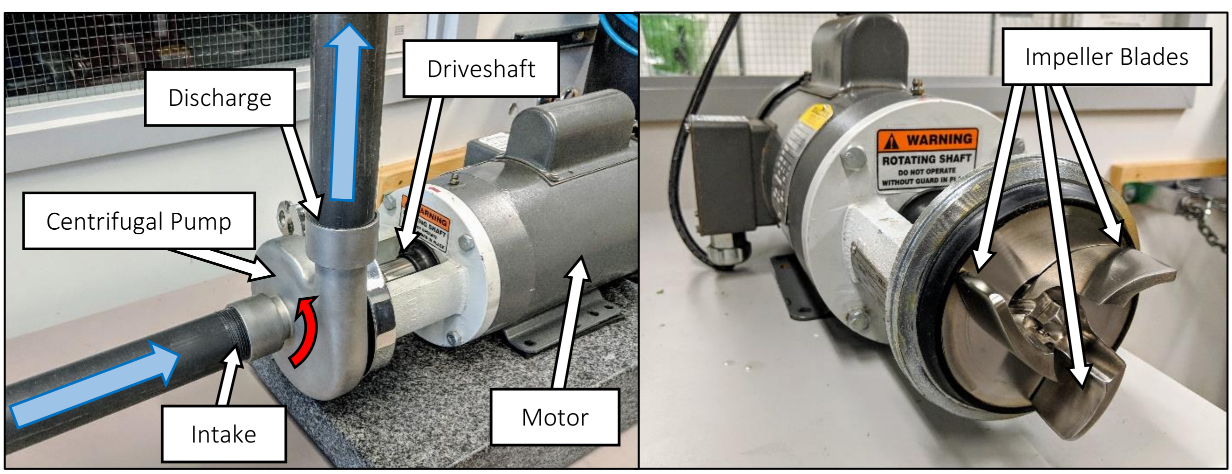


Figure 1.1: A 1.5 kW centrifugal pump and its components (left). Blue arrows indicate the direction of fluid flow. The red arrow specifies the direction of $\boldsymbol{\omega}$. Removing the housing reveals the impeller (right), which transfers kinetic energy to the working fluid.

As with any energy conversion process, there are losses. Understanding and minimizing these losses is central to the reliability, financial viability, and environmental sustainability of virtually every pumping process.

Centrifugal pumps are one of the world's primary consumers of electrical energy. In industries with extensive bulk fluid processing, centrifugal pumps can account for between 20 and 60% of expended electrical motor energy [2]. A 2001 study by the European Commission on improving the efficiency of pumps determined that centrifugal pumps were the largest single consumer of industrial electricity in the European Union [3]. The massive scale of electrical energy consumption by centrifugal pumps is not inherently negative, but instead underscores the potentially global impact that minor improvements (or declines) in pumping performance can have.

The causes and impacts of inefficient operation of centrifugal pumps have been extensively studied from a variety of perspectives [4]–[6]. However, despite this understanding, inefficiency is pervasive. A 2010 survey of pump manufacturers suggested that more than 40% of centrifugal pumps were incorrectly specified from the start and would suffer reduced efficiency throughout their existence as a result [2]. The European Commission efficiency study concluded, "[…] the largest energy savings are to be made through better design and control of pump systems." The study notes that inefficiently operating pumps are not strictly a consequence of inadequate system design, but also challenges in process management, such as timely access to technical expertise and inability to disrupt production to perform preventative maintenance [3]. The list could be expanded to include limited financial flexibility for process improvement and the tendency to oversize pumps to account for unanticipated changes in production volume. The underlying implication is that better performance cannot be achieved exclusively through better pump and pump control technologies. Improved performance is equally, or perhaps more dependent on how accessible and practical those technologies are to the end user.

The ability to quantify performance comprises not only the technical tools to make the measurements, but also the means to implement the tools in the first place, manage the data, and interpret the results. There are many commercial devices and measurement approaches for quantifying centrifugal pump performance, yet they are not widely implemented [7]. Understanding this phenomenon and giving consideration to the real-world barriers to implementing performance monitoring systems is paramount to creating a method that can traverse the divide between a novel research concept and an impactful industrial tool.

1.2 Pump Performance

The term "performance" can be defined in a variety of ways for centrifugal pumps. Informally, it is often employed as a vague synonym for efficiency, owing to the practical weight of minimizing waste energy in a given fluid process. However, more explicitly, performance can be considered as the conditions and phenomena that reflect the overall health of a pumping process. The most prominent elements include efficiency, operating cost, reliability, process stability, and the health of the mechanical components.

Depending on the type, the thermal efficiency of an exceptionally well-configured centrifugal pump can surpass 80%, meaning that, of the power delivered to the motor, 80% is realized as productive work on the operating fluid [8]. More conventional centrifugal pump efficiencies fall in the range of 30-70% [5], [9], [10]. Poorly controlled pumping processes may have efficiencies as low as 10%. The highest achievable efficiency for a given pump configuration is known as the Best Efficiency Point (BEP). Pumping efficiency, the BEP, and other performance factors are typically depicted in a characteristic *pump performance curve chart* produced by the manufacturer. The performance curve is a critical reference for centrifugal pump operators across industries. An example chart is shown in Figure 1.2.

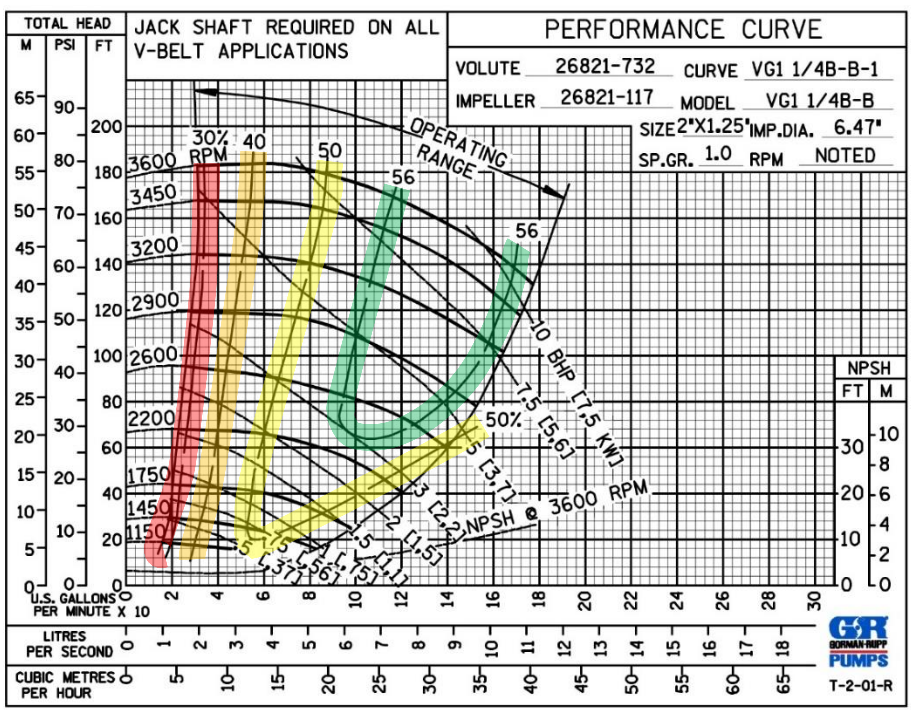


Figure 1.2: The performance curve chart for a Gorman-Rupp model VG1 centrifugal pump [11]. The highlighted curves are the efficiency profiles. Operating the pump at the flow (x-axis), pressure (y-axis), and impeller

speeds corresponding to the regions within these profiles will yield the denoted efficiency. For this pump, the BEP is 56%.

Centrifugal pumps degrade over time, which results in reduced efficiency and a departure from the BEP. The rate and severity of the performance decline can be a combined effect of a variety of factors; the fluid media, flow dynamics, impeller rotating speed, process uptime, operating temperature, maintenance, and more. Reductions in performance can be considered in two interrelated categories; losses due to mechanical wear and losses due to adverse fluid phenomena.

Mechanical wear performance losses are the result of structural changes to the pump components over time. These changes commonly include material erosion on the impeller surfaces, bearing degradation, corrosion, physical damage caused by process anomalies and flow debris, haphazard structural modifications to the impeller, or, perhaps most frequently, a combination thereof [7]. Performance losses due to wear are expected, even for a well-configured pumping process, but the impact can be exacerbated by poor management and lack of monitoring.

Losses from adverse fluid phenomena occur when a centrifugal pump expends energy on the operating media in ways other than productive fluid work. Typical scenarios include cavitation, gas entrainment, leaks, stalling, flow recirculation, and other volumetric losses [12]. Adverse fluid phenomena can cause, or be caused by, mechanical wear, improper flow configuration, flow anomalies, and process control problems.

If the mechanical degradation or adverse flow conditions are severe enough that the pumping process is profoundly altered, a pump's behavior can depart from the original performance curve entirely, leading to an operating state without an established reference for performance. In industry, pumps operating in reference-less states are common and notoriously difficult to manage [13]. When the performance curve of a centrifugal pump is rendered unusable, whether by mechanical wear or flow irregularities, additional means to measure and monitor performance are vital.

1.3 Performance-monitoring Systems

Just as there are many phenomena that comprise a centrifugal pump's overall performance, there are equally many tools and methodologies for measuring it. Before engaging in a discussion of the present technical avenues for quantifying centrifugal pump performance, it must be noted that performance-monitoring should not be considered exclusively from the domain of the end-user (though, being that the end-user is typically the person implementing the system, it often is). The previous section notes industry's strong reliance on the manufacturer's performance curve chart and the difficulty that can arise when that reference no longer suits a particular centrifugal pump. The intuitive next step is to jump to consideration of secondary monitoring tools that can be used to supplant the lost performance reference. However, a more thoughtful approach will consider both the end-user and the pump manufacturer. This perspective expands the applied problem from an assessment of discrete performance-monitoring devices to a broader discussion of how pump performance data is collected and what additional references would improve its value. More succinctly, if the pump manufacturer provided the end-user with something more robust and flexible than a static performance curve chart, what kind of impact could that have on the efficacy and popularity of performance-monitoring systems? As data analytics and machine learning become more widely accessible, the question becomes more intriguing.

Efficiently controlling a centrifugal pump relies heavily on the operator's means to understand the phenomena at hand. To a degree, this understanding is based on expertise with the specific system and fluid process. Centrifugal pump operators can often identify adverse conditions simply by listening to a pump, or consulting basic temperature or pressure measurements if they are obtainable [7], [13]. However, in the case of slowly progressing conditions or unfamiliar fluid phenomena, subjective observation becomes problematic. In these cases, unless more accurate performance-monitoring tools are used, inefficient operating states can persist indefinitely. This underscores the importance of having the ability to dependably measure and quantify centrifugal pump performance.

Despite their relative lack of widespread use, there are an abundance of sensor systems and monitoring approaches for observing the performance of centrifugal pumps. The following sections summarize the operating bases of the foremost commercial methods and associated research. Commercial products typically fall into four groups: conventional sensors for process measurement, efficiency monitoring devices, vibration analysis systems, and fault detection systems based on machine learning. Beyond the commercial domain there are experimental sensors and methodologies that have been demonstrated in limited applications but have yet to reach wide commercialization. These include magnetic wear sensing, hydrophone monitoring, motor phase current sensing, and dynamic pressure analysis.

1.3.1 Conventional Sensors

Though many industrial centrifugal pumps operate without instrumentation [7], it is common to use conventional sensors to quantify basic elements of a pump's operating state. Direct measurements of fundamental process parameters can be used to infer the performance of centrifugal pumps without additional analysis or signal processing. Widely used sensors include pressure transducers, thermometers, flow meters, tachometers, and wattmeters (Figure 1.3). Implementations of conventional pump process monitoring can include intake and discharge pressure measurements, bearing temperature readings, flow rate monitoring, and rotational speed measurement.

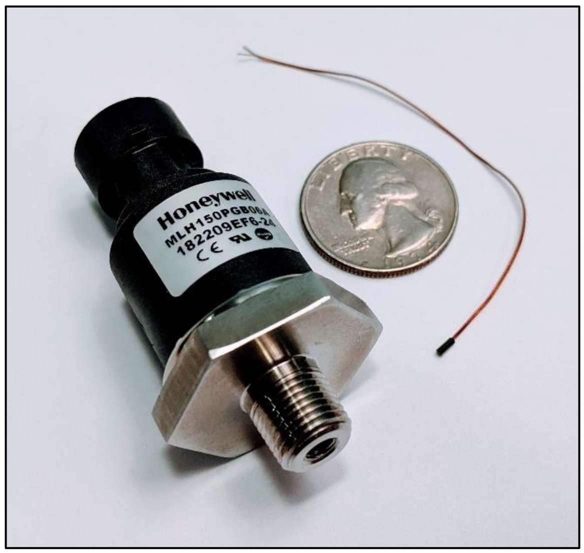


Figure 1.3: A piezoresistive pressure transducer (left) and a negative temperature coefficient (NTC) thermistor (right). Pressure and temperature sensing are two of the most common methods to monitor pump operation.

Conventional sensing is generally applied as real-time, steady-state process measurements. This approach has the benefit of being affordable, reliable, and straightforward to implement and interpret. When a drastic variation occurs, or a measurement deviates from a required threshold, it suggests to the operator that some element of the pump's operation has changed and needs attention. However, this reactive approach is in itself insufficient for classifying complex or slowly progressing adverse conditions. Without a broader system for recording data, analyzing trends, and validating phenomena, conventional process measurement falls short in the depth of information it can provide.

1.3.2 Efficiency Monitoring

Efficiency-based performance-monitoring systems provide the pump user with a measurement of how effectively a centrifugal pump is converting input energy into productive fluid work. There are two main methodologies to quantify pump efficiency; conventional thermal efficiency and the thermodynamic efficiency method.

The conventional thermal efficiency method correlates the pump's fluid power output W_{fluid} to the input power W_{in}. The ratio

$$\eta_{\mathrm{conv}} = \frac{W_{\mathrm{fluid}}}{W_{\mathrm{in}}} \tag{1}$$

defines the thermal efficiency of the pumping process. In the case of a centrifugal pump, $W_{\mathrm{fluid}} = \rho g H Q$, where ρ is the average fluid density, g is the acceleration of gravity, H is the average pressure head across the pump, and Q is the average volumetric flow rate through the system. Equivalently,

$$W_{\mathrm{fluid}} = \Delta p Q\ , \tag{2}$$

where Δp is the difference between the average pressures at the pump intake $\overline{p_{\mathrm{in}}}$ and discharge $\overline{p_{\mathrm{out}}}$. Substituting (2) into (1) yields the conventional determination for a centrifugal pump's thermal efficiency,

$$\eta_{\mathrm{conv}} = \frac{\Delta p Q}{W_{\mathrm{in}}}\,. \tag{3}$$

This efficiency determination method requires four sensors to implement; pressure sensors at the pump intake and discharge to measure Δp, a flow meter to determine Q, and a wattmeter to measure W_{in}. Where performance monitoring is used, this approach is common. It is the method specified by ISO 9906:2012 for hydraulic performance testing of rotodynamic pumps [14], and specific implementations have been proposed by many researchers [10], [15]. It is also used in commercial devices [16], [17].

For centrifugal pumps generating more than 1000 kPa, or pumping processes where attaining an accurate flow rate or power measurement is impracticable, it may be more suitable to use a thermodynamic approach and determine efficiency through direct measurement of waste energy, rather than as a ratio of output and input work [18]. The thermodynamic efficiency

$$\eta_{\mathrm{therm}} = \frac{W_{\mathrm{i}}}{W_{\mathrm{in}}}\,, \tag{4}$$

where W_{i} is the power required to operate an ideal pump. In an isentropic compression cycle of a fluid with mass flow rate $\dot{m}$ and average specific volume $\bar{v}$ from p_{in} to p_{out}, the input power

$$W_{\mathrm{i}} = \dot{m}\int_{p_{\mathrm{in}}}^{p_{\mathrm{out}}} v dp = \dot{m}\bar{v}\Delta p = \dot{m}\frac{1}{\rho}\Delta p\ . \tag{5}$$

Even in an ideal pump, there is a small temperature increase when the liquid is compressed. If it can be reasonably assumed that heat loss from the pump to the atmosphere is negligible, temperature difference between the pump discharge to intake $\Delta T = T_{\mathrm{out}} - T_{\mathrm{in}}$ can be defined as

$$\Delta T\ = \Delta T_{\mathrm{i}} + \Delta T'\ , \tag{6}$$

where ΔT_{i} is the expected temperature increase from ideal compression and $\Delta T'$ is the additional temperature rise resulting from pumping inefficiency. The value of ΔT_{i} is determined from established property tables for the working fluid undergoing compression $\Delta \mathrm{p}$. Using the specific heat at constant discharge pressure C_{po}, wasted power can be equated to the increase in enthalpy beyond the ideal compression, such that

$$W_{\mathrm{in}} - W_{\mathrm{i}}\ = \dot{m}C_{po}\Delta T'\ . \tag{7}$$

Combining (4) through (7), the thermodynamic efficiency determination

$$\eta_{\mathrm{therm}} = \frac{1}{1+\rho C_{po}(\Delta T - \Delta T_{\mathrm{i}})/(\Delta \mathrm{p})}\ . \tag{8}$$

The thermodynamic approach to efficiency measurement also requires four sensors; pressure transducers at the pump intake and discharge to measure Δp, and temperature sensors at the intake and discharge to determine ΔT. It has a benefit in that it eliminates the need for flow rate and power sensing, which can be costly to integrate. Applied in accordance with their respective constraints and assumptions, each efficiency measurement approach will yield an equivalent result. This sensing technique is also prevalent in both commercial monitoring and academia [19].

Each efficiency determination approach has disadvantages. For the conventional thermal efficiency determination method, the main criticism is in the need to measure flow rate and input power. A. Whillier highlights the common case in chemical transport where a pump efficiency measurement is desired, but the diversity of working fluids makes reliable flow rate measurement with a single sensor unattainable [20]. Similarly, A. Cattaert describes the challenges in accurately determining input power in mining applications where a pump is not powered by an electric motor [18]. In addition, first-hand discussions with pulp and paper producers have made clear the economic infeasibility of integrating flow and power sensors when a single mill may have hundreds of unique centrifugal pumps in operation [13].

The thermodynamic efficiency determination method, despite the simplicity of the sensors involved, also has drawbacks. The main concern is the accuracy by which ΔT can be measured. To satisfy its highest precision class, ISO 5198: Centrifugal, Mixed Flow, and Axial Pumps Code for Hydraulic Performance Tests specifies this method as only suitable for centrifugal pumps generating at least 100m head (980 kPa) [21]. In his study of the required measurement uncertainties for this efficiency determination approach, A. Milne finds that a low pressure centrifugal pump generating 200 kPa would require temperature measurement within 4 mK to determine the efficiency within 5% – not reasonably achievable in a typical industrial setting [22].

In terms of broader performance classification, efficiency-based monitoring falls short from a variety of perspectives. At a fundamental level, meaningfully interpreting an efficiency measurement requires an understanding of the expected efficiency. In an application with hundreds of pumps in varying configurations, the challenge of maintaining a suitable reference for good efficiency for each is significant. On the inverse, even when an inefficient operating state is truthfully detected, the cause of the loss cannot necessarily be extrapolated from the efficiency determination itself. This diminishes the effectiveness of standalone efficiency monitoring in remedying unhealthy operating states. A measure of efficiency is useful to have, but requires additional data if it is to be used to make substantive judgments about a centrifugal pumping process.

Finally, from a process management standpoint, the intent of measuring efficiency is to avoid inefficient operating conditions. Yet, this approach to performance monitoring requires the pump to have *already* reached a degraded state before the inefficiency can be recognized. This can be mitigated somewhat by establishing efficiency trends and preventative maintenance schedules, but maintaining such a record introduces its own practical challenges. In addition, characterizing a pump's performance exclusively on efficiency limits the operator's ability to recognize serious conditions that may not correlate strongly to efficiency losses, such as minor cavitation or flow debris. These shortcomings ultimately discourage the widespread use of standalone efficiency monitoring systems.

1.3.3 Vibration Monitoring

Vibration monitoring is the application of accelerometers for the purpose of evaluating characteristic vibrations that propagate through the pump's mechanical structure. In a typical implementation, accelerometer data is collected from an external surface on the pump[1] and converted to a frequency spectrum using fast Fourier transform (FFT). The frequency features can be manually interrogated and correlated to known pump phenomena or integrated as parameters into a broader diagnostic system or machine learning algorithm. An example of the frequency content of a centrifugal pump accelerometer measurement is shown in Figure 1.4.

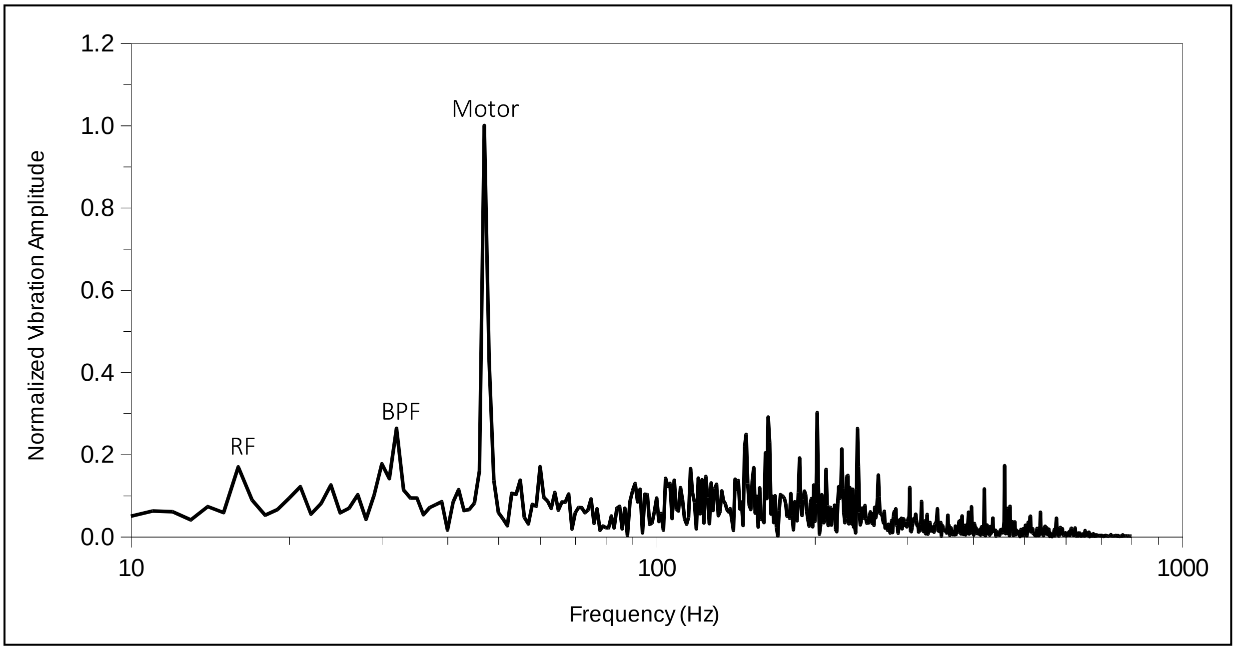


Figure 1.4: The normalized frequency spectrum of a 30 kW centrifugal pump with a two-blade impeller, collected using a single-axis accelerometer on the axial face of the pump volute. Here, the pump is operating at a rotating frequency (RF) of 15.4 Hz, which generates a small peak at the corresponding frequency. Likewise, the blade-passing frequency (BPF), equal to twice the RF, exhibits a modest peak. However, the spectrum is dominated by a vibration mode occurring at three times the RF (47.3 Hz). This is the undesirable influence from the close-coupled driving motor, which uses three-phase power delivery, transmitting vibrations through the driveshaft into the pump's resonant structure. Frequency content beyond 800 Hz has been removed using a low-pass filter.

[1] It is noted that airborne acoustic measurement is occasionally used in lieu of structure-borne vibration monitoring of centrifugal pumps. In terms of analysis, the signals and features can be treated similarly. However, when conducting acoustic emission measurements using a microphone, the influence of environmental sound can be a major detriment. In many noise-filled industrial settings, acoustic pump monitoring is not practicable. As such, it is not explicitly considered in this work.

The mechanical vibration characteristics of centrifugal pumps have been extensively studied [23]–[25]. Vibration monitoring in the frequency domain is particularly suited for distinguishing cyclic mechanical pump phenomena [26] and motor faults [27]. Behaviors such as a warped input shaft, bearing flat spots, lack of lubrication, and unbalanced components have distinct spectral manifestations that can be classified by a trained operator. C. Scheffer characterizes sources and diagnosis methods for a variety of mechanical-fault-induced centrifugal pump vibrations with consideration to informing remedial actions and preventing downtime [28]. He also discusses the correlation between accelerometer placement and how prominently a particular behavior manifests in the acquired measurement. M. Abd-Elaal *et al.* apply vibration monitoring to identify mechanical irregularities in a small centrifugal pump [29]. They demonstrate shaft imbalance, impeller damage, and pedestal looseness using features in the vibration frequency spectrum.

The characteristic vibrations of an operating centrifugal pump can also contain evidence of fluid phenomena. A. Abdulaziz *et al.* investigate the frequency spectrum of vibration measurements to determine the presence and severity of cavitation [30]. They establish that the inception of cavitation can be detected through discrete vibration levels corresponding to the rotating frequency (RF) and blade passing frequency (BPF). The BPF is equal to the RF multiplied by N_{imp}. Specifically, as cavitation evolves, the energy peak at the BPF weakens dramatically, and noise beyond twice the BPF becomes the dominant spectral feature.

An evident shortcoming of vibration-based monitoring systems is that some phenomena simply do not propagate vibrations strongly enough to be measured externally [31]. In the case of mechanical faults, despite the direct transmission path through the pump structure and into the accelerometer, the observable features in a particular sample are heavily dependent on physical placement of the sensor [38]. For fluid-borne phenomena, where the vibrational energy must first transfer from the fluid to the pump structure, the vibration transmission may be insufficient to make a reliable measurement. Conversely, some sources of vibration can propagate disproportionately strongly through the pump structure, obscuring more subtle phenomena. This is demonstrated in Figure 1.4, where vibrations caused by power delivery to the electric motor have a detrimental influence on the vibration spectrum. Environmental structural noise generated by large equipment vibrating nearby could yield a similar outcome. In these applications, vibration monitoring would lose its efficacy.

Another concern with vibration analysis is the relative lack of generalizability of accelerometer measurements between pumps. Specifically, because each individual centrifugal pump is a unique resonating structure, the vibrational energy transferred from a particular phenomenon to the transducer may not be consistent. The generalizability problem worsens when considering pumps of dramatically different sizes and forms. While the trends may be shared, the amplitudes, frequencies, and relative manifestations of performance features will likely be considerably different.

1.3.4 Machine Learning for Performance Monitoring and Fault Detection

The scope of insights gained from centrifugal pump performance monitoring can be augmented by using the collected information as inputs for machine learning algorithms. Commercial fault detection systems generally rely on vibration measurements, but multi-sensor learning models have also been demonstrated.

To classify flow blockages and impending cavitation, A. Panda *et al.* implement a Support Vector Machine (SVM) algorithm on a series of statistical features of vibration measurements taken on a centrifugal pump's casing and bearing housing [32]. Using a feature vector comprised of mean, standard deviation, kurtosis, crest factor, and signal entropy, they develop and test binary and multi-class classifiers that aim to detect a) flow conditions ranging from unobstructed to 67% blockage across a range of pump rotating frequencies and b) normal flow versus an impending cavitation state (which is designated as just prior to the formation of visible vapor bubbles in the pump discharge). The flow obstruction classifier demonstrates test accuracies ranging from 61.4% (achieved in the lowest speed, lowest blockage percentage scenario) to 100% (achieved in the highest speed, highest blockage percentage scenario), with most classification accuracies falling between 80 and 100%. The cavitation classifier demonstrates near-perfect accuracy across operating speeds. While the experiment yields a high classification accuracy across the sampled states, the feasibility of using this approach to detect these conditions in industry is low. Flow obstructions, particularly significant ones, where the learning algorithm is shown to be most accurate, would be readily detected through conventional measurement with a pressure transducer. In addition, regarding the classification of healthy or near-cavitating states, the conditions can be distinguished by manual observation of the frequency spectrum or standard deviation, skewness, or kurtosis trends. As such, employing a machine learning algorithm to achieve classification of these extreme states is not particularly advantageous.

V. Muralidharan *et al.* compare a series of discrete wavelets of accelerometer measurements using a decision tree algorithm for the purpose of identifying the most suitable signal features to employ in classification of bearing faults, impeller damage, and combined scenarios [33]. Using the Reverse

Biorthogonal 1.5 (rbio1.5) wavelet, they are able to classify the mechanical condition with 99.84% accuracy. The same authors compare naïve Bayes and Bayes network classifier algorithms, using wavelets, for the same purpose and achieve similarly high classification accuracy [34]. However, as with the study by A. Panda *et al.*, the phenomena being identified would likely be conspicuous through manual observation and would be ascertained by more direct methods in practice. Direct classification of comparable faults through frequency analysis has been experimentally demonstrated by A. Daraz *et al.* (via acoustic emission measurements) among many others and analytically by M. Zhang *et al.* [35], [36]. It is a core function of most commercial vibration-based centrifugal pump fault detection systems [37], [38].

However, there are cases when mechanical and fluid phenomena cannot be readily classified through manual data analysis and machine learning is necessary. M. Shervani-Tabar *et al.* employ a multi-class SVM algorithm for the purpose of characterizing the severity of cavitation in an axial-flow pump using vibration measurements [39]. Whereas the presence of cavitation is typically observable through conventional monitoring methods, the nature of the severity is less so. Their work suggests that machine learning methods may have particular value in assessing the severity of pump phenomena, rather than simply the presence or absence.

1.3.5 Alternative Sensors and Monitoring

There are elements of centrifugal pump performance that cannot be easily characterized by conventional process measurement, efficiency monitoring, vibration-based analysis, or associated machine learning methods. A ubiquitous example is the state of mechanical erosion on the impeller. Several purpose-built sensors and methods have been developed for this case, yet none have achieved wide-scale implementation [7]. An externally mounted sensor that measures the extent of material loss from a centrifugal pump impeller has been developed [40]. The device uses an inductive coil to drive magnetic flux through the operating area of a centrifugal pump impeller. As material erodes from the impeller with time, the fluid gap between the impeller and pump side plate grows. This increasing gap raises the reluctance of the magnetic circuit, which is then measured and correlated to determine the extent of impeller material loss within 0.25 mm. The sensor has been demonstrated in a controlled trial, however its adaptability to pumps of various sizes is limited by its physical dimensions. Additionally, the high-magnetic-permeability material of sensor's flux guide is relatively expensive, difficult to machine, and fragile, making it suitable for a demonstration of the concept, but not industrial implementation. Significant refinement would be needed to employ the sensor on a wide scale.

T. Ahonen *et al.* demonstrate an unconventional monitoring approach that applies phase current measurements to extrapolate shaft power, flow rate, and efficiency of fixed-speed centrifugal pumps [41]. The method is intended to provide a rapid, non-invasive, coarse means of auditing a pump's operating state, which would be used to inform the user if more thorough monitoring is warranted. With a known nominal rotating speed, their method experimentally determines shaft power and flow rate within 3% and 16%, respectively. In conjunction with a performance curve chart, these measurements can allow for an efficiency determination. Most critically, this approach is only suitable for an approximation. Its intended purpose is to alert the pump operator of severe performance deterioration, such as running dry or stalling. The method is not suitable for characterizing less extreme adverse operating conditions. In addition, it relies on an accurate performance curve chart to make efficiency determinations. As discussed previously, this reference is frequently unavailable to the pump operator. Finally, the method has no means to capture types of wear that do not have a conspicuous effect on the pump shaft power, such as misalignment, bearing wear, or minor entrainment of gases.

S. Yuan *et al.* experimentally contrast the application of hydrophones, accelerometers, and dynamic pressure transducers to monitor pump flow noise phenomena in a healthy centrifugal pump [42]. They distinguish the pump phenomena into two groups; discrete frequency noise caused by cyclic behaviors and broadband noise caused by chaotic fluid phenomena. The frequency peaks of the resulting spectra from each of the three sensing methods are strongly correlated, but with varying magnitudes, likely due to the sensitivity and response of the respective sensors utilized, as well as the resonant behavior of the testbed pump. Similar results are reported by A. Suhane [43]. Yuan successfully shows that each sensing approach can be used to monitor flow noise, but the presence of normal structural noise (from the motor, bearings, etc…) impacts each measurement differently, clouding the analysis. The comparison also does not evaluate which method is most suitable for a given fluid phenomenon. In addition, the study utilizes a pressure transducer with a rise time of 1 μs – nearly three orders of magnitude faster than a conventional liquid water pressure transducer [44], [45], making it operate, effectively, as a wall-mounted hydrophone.

Though uncommon commercially, dynamic pressure monitoring has also been demonstrated experimentally. X. Li *et al.* evaluate the dynamic pressure signals of a centrifugal pump to characterize the onset of cavitation [46]. The relationship between cavitation and pressure fluctuations is captured by interrogating the probability density function (PDF) of the resulting pressure signal. Observing the peak of the PDF as a function of the cavitation severity, which is determined by the available net positive suction head (NPSHA), the study shows a dramatic increase as cavitation occurs. This validates that dynamic pressure monitoring is an effective way to observe cavitation, but as discussed previously, cavitation is generally a severe phenomenon that can be captured using a variety of less intrusive methods. E. Higham *et al.* perform a related experiment, using direct evaluation of the pressure transducer frequency content in their analysis [47]. In addition to cavitation, the Higham study characterizes flow obstructions and running-dry operating conditions. However, as with cavitation, all the conditions evaluated are severe and readily observable through a variety of methods.

1.4 Phenomena of Interest

Centrifugal pump performance is impacted by a variety of dynamic phenomena, many of which are well understood in both the academic and applied senses. These include the mechanical issues like lack of lubrication, bearing wear, misalignment, impeller damage, acoustic emission, and shaft play, as well as fluid-borne behaviors like cavitation, recirculation, flow debris, flow obstruction, turbulence, and hydraulic shock. The list could tangentially include phenomena related to the electric motor and power transmission as well. The body of study on these behaviors is extensive, therefore they are not further discussed in this research. Instead, this work explores two centrifugal pump phenomena that are persistently difficult to measure and characterize in a typical industrial setting; flow entrainment of dispersed gas bubbles and radial impeller wear. This section describes the physical behavior and associated measurement techniques for each.

1.4.1 Gas Entrainment

The centrifugal pumps considered for this study are designed for transporting only liquids. Gas entrainment is the undesirable presence of two-phase gas/liquid flow entering the centrifugal pump (Figure 1.5).

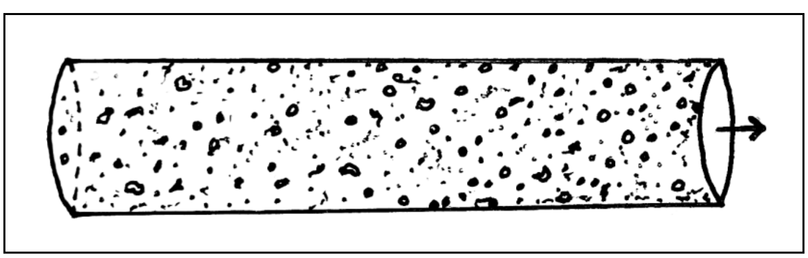


Figure 1.5: Distributed gas bubbles entrained in liquid flow through a pipe.

There are two[2] primary modes of gas entrainment: dispersed two-phase flow (i.e. "bubbly") and swirling flow with a separated gas phase in the center ("vortexing"). Vortexing is common in fluid systems where a pump pulls fluid from an insufficiently full reservoir, generating a column of air [49]. Distributed two-phase flow can occur by a variety of mechanisms, including intake flow obstructions (where the pressure at the pump intake is reduced and a suction created, turning vents into inlets for air) or a reservoir that contains a two-phase mixture from a previous process that has had inadequate time to separate before being pulled into the pump [9].

The mechanics and performance impacts of vortexing two-phase flow in centrifugal pumps have been characterized by P. Bardet *et al.* [50] and T. Schäfer *et al.* [51]. F. Gülich notes that the performance detriments and physical manifestations of the ingress of a separated column of entrained air into a centrifugal pump are conspicuous [9]. It is therefore considered to be a condition that can be sufficiently characterized with current performance monitoring approaches. This research focuses instead on dispersed two-phase conditions.

T. Schäfer *et al.* experimentally demonstrate the performance impacts of varying volumes of gas entrainment in a centrifugal pump moving liquid water [51]. Performance is quantified using the hydraulic power output from the pump, normalized to a healthy (non-entrainment) condition. For the distributed two-phase case, the study demonstrates an approximately linear decrease in relative hydraulic power output from the healthy state to an entrained air void fraction of 5%, at which point the power has dropped to 25% of normal. Additionally, they employ high-resolution gamma ray computed tomography to quantify the phase fractions and gas distribution patterns (i.e. holdup patterns) on the impeller surface. The study demonstrates that gas entrainment can cause a major detriment to performance, but the reported trends apply to a singular pump configuration, operating speed, and pressure head. It can be assumed that the impacts of gas entrainment would vary with configuration changes. In pump configurations where the decrease is less severe (or not measurable), fluid power would not be suitable feature for characterizing the volume fraction of air. In addition, to draw more thorough performance conclusions from the trends, a broader set of operating speeds and pressures is needed.

---

[2] In related research, more precise descriptors of dispersed gas entrainment conditions are sometimes necessary. T. Xie *et al.* use the following terms, listed from the most distributed to most consolidated bubble formation: dispersed bubbly flow, layered bubbly flow, incipient plug flow, plug flow, churn-turbulent flow, and slug flow [48]. While each can be classified as a unique phenomenon, from the application perspective of this research, the distinction is not strictly necessary.

M. Stan *et al.* analytically and experimentally explore the impacts of gas entrainment on a multi-stage centrifugal pump's vibration frequency spectrum, acoustic emission, and efficiency [52]. The study evaluates dispersed air-water mixtures between 0% and 11% volume fractions (normal and stalling conditions, respectively). They demonstrate that the introduction of air into the fluid flow changes the pump's characteristic performance curve, which implies that presence of two-phase flow can render the manufacturer's performance curve chart inaccurate. The study also shows that at higher flow rates, the respective change in vibration spectral amplitude, caused by increasing the void fraction of air in the fluid, diminishes. This suggests that accelerometer measurement may not be sufficient for observing the severity of gas entrainment in some operating conditions.

Dynamic pressure sensing has been successfully applied to investigate multi-phase flows. Q. Si *et al.* explore the frequency content resulting from pressure pulsations caused by varying degrees of air entrainment through two different pumps [53]. They report an increase in broadband low frequency noise in the fluid pressure fluctuations as the volume fraction of air increases. Comparable spectral changes are reported by Y. Xu *et al.* [54]. The measured spectra do not vary significantly as the transducer position changes within the pump volute. They also demonstrate a decrease in the frequency-domain energy peak corresponding to the BPF. Si shows that dynamic pressure sensing has the potential to adequately capture the effects of gas entrainment, but significance of the associated statistical features of the signals must be more thoroughly characterized if this sensing method is to be used in the application of performance monitoring.

T. Xie *et al.* apply dynamic pressure sensing in conjunction with an artificial neural network (ANN) to classify the bubble morphology of entrained air in three-phase water-air-pulp flows in two related studies [48], [55]. In the first experiment, three pressure sensors are located in steady flow downstream of an injection nozzle. The transducers are placed sufficiently far from the pump discharge such that the dynamic pressure phenomena caused by the pump itself are not present. A multi-layer perceptron (MLP) with a single, seven-node hidden layer is constructed using standard deviation, skewness, kurtosis, and several time-shift autocorrelation values from each pressure sensor as input features. The outputs are classified into "bubbly," "chug," "churn," and "slug" flow regimes. The volume fraction of entrained air is reported only as "not more than a few percent." Experiments are performed at 0.5, 1.0, and 1.5% pulp suspension consistencies, for a total of 197 recorded states. Using this method, they successfully classify the type of bubble formation in 90% of the training cases. In the second study, they perform a similar experiment, instead using a single pressure sensor. From the fluctuating pressure measurements, they compute: a) the sum spectral energies in a range of frequency bands (0-3, 3-8, 8-13, 13-25, and 25-30 Hz), b) the mean frequency (i.e. spectral centroid), and c) variance. These measures form the input features for a multi-class MLP. In this configuration, they achieve correct bubble regime classification with 87% accuracy.

The studies by T. Xie demonstrate the potential for using dynamic pressure sensing in conjunction with neural networks for classifying air entrainment. However, these studies explore the *morphology* of air-entrainment, not the severity. The experiments are also conducted on three-phase water-air-pulp mixtures, whose dynamic fluid behavior does not necessarily correlate to two-phase air-water mixtures or other fluid combinations. Further study is required to determine if this method is suitable for a sufficiently wide scope of volume fractions of air and constituent fluids. Additionally, Xie's experiments do not consider the potentially dramatic dynamic influence of the pressure fluctuations caused by the centrifugal pump. To apply either of the reported methods for classification of the severity of gas-entrainment in pumps, more research is necessary.

1.4.2 Impeller Wear

Impeller wear refers to undesirable mechanical erosion from the surfaces of a centrifugal pump's impeller (Figure 1.6). It is a natural occurrence in pumps, but the rate of degradation can be aggravated by poor process maintenance. Impeller wear limits the lifetime of a centrifugal pump and can cause inefficiency and serious damage if not addressed appropriately [56].

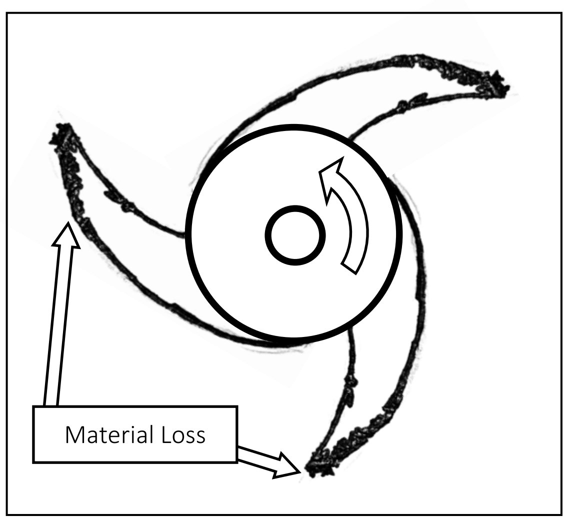


Figure 1.6: Erosive wear on the tips of a three-blade impeller.

S. Krüger *et al.* differentiate the causes of impeller material loss into two categories; shock-type processes, where particulate matter in the flow impacts the impeller surface, and friction-type processes, where adverse flow patterns and turbulence cause erosion over time [6]. They describe shock-type wear as most prevalent in liquid/solid multi-phase pumping applications, such as in mining, paper production, and marine applications, where pumps are expected to move liquids with abrasive constituents. The study shows that shock-type erosion patterns manifest most severely on the leading surface of the impeller, near the fluid intake. Similar results are demonstrated by Y. Wang *et al.* [57] and A. Daraz *et al.* [35]. In the case of friction-type wear processes, which are present in all centrifugal pumps, regardless of the flow media, Krüger demonstrates that the greatest material loss occurs at the trailing tips of the impeller blades, which shrinks the effective diameter of the impeller, and at the impeller surface abutting the side plate, which can increase the incidence of recirculation within the pump. Their work suggests that both shock-type and friction-type wear processes can degrade the performance of the centrifugal pump.

The primary focus of this thesis is in classifying erosion resulting from friction-type wear, due to its applicability to all centrifugal pumps, rather than only pumps transporting solid content. Specifically, radial material loss from the impeller tips is considered, as it is experienced by all centrifugal pumps, regardless of impeller configuration.[3] The significance is demonstrated by M. Matlaka *et al.*, who use numerical methods to characterize the flow performance impacts resulting from material removal from the tips of impeller blades [58]. They demonstrate a 40% decline in pumping power after reducing the impeller diameter by 15%, suggesting that characterizing impeller tip erosion has fundamental importance in evaluating the operating health of a centrifugal pump.

A. Suhane experimentally characterizes the changes in pressure fluctuations, structural vibration, and airborne noise resulting from increasing the clearance between a rotating impeller and stationary diffuser vanes [43]. The study employs a 7.5 kW centrifugal pump, operating at 1440 RPM. Dynamic measurements are interpreted in the frequency domain using FFT. For the pressure samples, the author demonstrates that increasing the nominal clearance between the impeller and stationary vanes by a factor

[3] Wear along the side plate surface occurs primarily in centrifugal pumps with open-type impellers. Impellers come in "open," "semi-open," and "closed" configurations. Open impellers have no shrouds on either face of the blades, and must maintain a close gap with the side plate to avoid recirculating flow losses. This is the type of impeller wear investigated in [40]. Semi-open impellers have an integrated shroud on the face opposite the pump intake. Closed impellers have shrouds on both faces of the impeller, turning the gaps between the blades into flow channels. On pumps that use semi-open and closed impellers, this wearing surface is not present.

of 4.5 (1.5 mm to 6.8 mm) results in an approximate 50% amplitude decrease at the BPF. Reductions in the amplitudes of the associated frequency in the structural vibration and airborne noise measurements are also observed, but to a lesser extent. Frequency impacts outside the BPF and associated harmonics are not reported. Though the study specifically explores the spectral impacts of clearance between a pump impeller and diffuser vanes, the experimental configuration is analogous to the growing clearance between a radially wearing impeller and the pump's discharge port. It suggests a correlation between erosion of the impeller diameter and the pump's vibration, in both the fluid and structural sense. Further, Suhane shows that the relationship is most prominent when observed as fluid pressure fluctuations, rather than through structural vibration or emitted noise. Further study is needed to clarify this correlation.

A. Jami *et al.* employ a machine learning approach to classify a damaged impeller in a centrifugal pump based on accelerometer measurements [31]. They apply a MLP with binary classification to identify the presence of impeller cracking and unbalance using an accelerometer. The machine learning method uses a feature space consisting of time-domain features, including kurtosis, root-mean-square (RMS), skewness, and variance, as well as frequency-domain features, including spectral frequency peaks and wavelet decomposition coefficients. The MLP contains one hidden layer and achieves near perfect classification across the majority of the trials when using the entire feature set. However, the experimental impeller exhibits wear that would be considered exceptionally severe in practice. The peak MLP performances based strictly on time-domain or frequency-domain features are less exceptional, reported as 63% and 74%, respectively. The authors report difficulty in identifying slight changes in impeller condition using this method, particularly when only considering time-domain statistical features. This reasons one of two conclusions: that time-domain statistical features may not contain satisfactory evidence to identify minor changes in impeller wear, or that accelerometer measurement may not be the most suitable method for capturing subtle changes in the impeller structure. The work of Suhane discussed above [43] suggests the latter, though further substantiation is necessary.

L. Cao *et al.* conduct a numerical investigation of changes in fluid pressure fluctuations resulting from subtle changes in the axial gap between an impeller and pump housing, consistent with an impeller wearing against its side plate [59]. They apply the k-ω turbulence modeling method, which they argue is most suitable for prediction of rotating, separating flow. The study shows that pressure fluctuations in general are most significant near the volute discharge, decreasing with distance down the discharge pipe. Cao confirms that increasing the axial impeller gap by as little as 1 mm results in observable changes in lower frequency pressure fluctuations. The study does not show dramatic change in the energy peaks at the dominant frequencies, but this is to be expected when considering minor axial wear, which contributes to turbulence and recirculation, but would not change the pressure fluctuation modes associated with the impeller radius. Their work suggests that dynamic pressure measurement is suited for the subtle fluid phenomena associated with minor wear. Additional study is needed to demonstrate that this modeling method is applicable for pressure fluctuations associated with radial, rather than axial wear.

1.5 Pressure Dynamics in a Rotating Pump

A schematic highlighting the terms and positions of the primary mechanical elements of a centrifugal pump is shown in Figure 1.7.

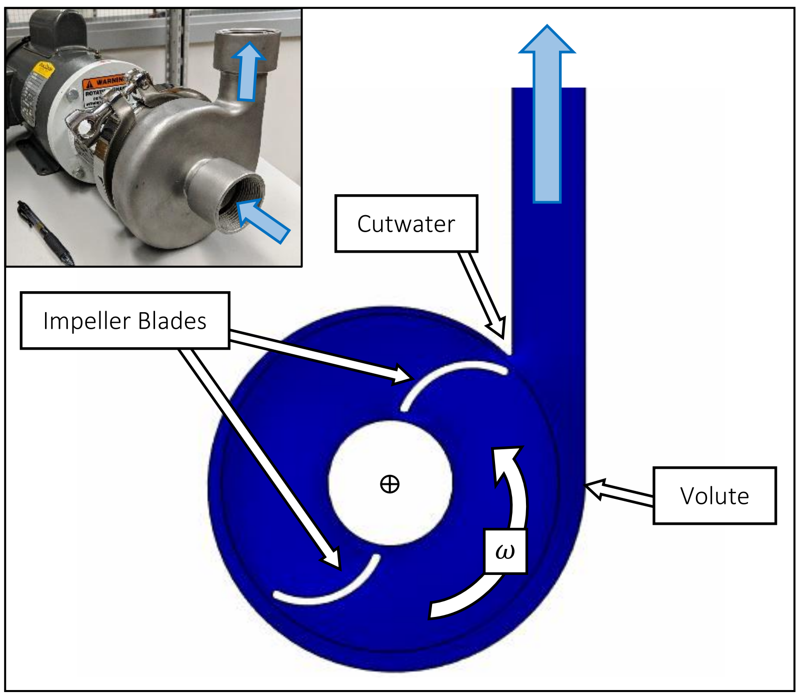


Figure 1.7: A two-dimensional schematic of a simplified centrifugal pump with a two-blade ($\boldsymbol{N_{\mathrm{imp}} = 2}$) impeller. The inset shows a perspective view of a comparable physical pump for clarity. Fluid flow direction is indicated by the blue arrows. The operating fluid enters the pump through the port at the center of the impeller (from out of plane) and exits vertically. A scroll-shaped funnel called the volute houses the impeller and guides the working fluid from the area of the rotating impeller to the discharge port. The notch at intersection of the curves of the volute scroll is the "cutwater."

1.5.1 Ordinary Operation

During normal operation, fluid enters the pump body through the intake port with pressure p_{in}, average velocity $\overline{v_{\text{in}}}$, and temperature T_{in}. Fluid exits through the discharge port with pressure p_{out}, average velocity $\overline{v_{\text{out}}}$, and temperature T_{out}. When a centrifugal pump moving nominally incompressible fluid has reached steady state operation, Q is constant. Consequently, in the case where the cross-sections of the pump intake and discharge are equal in area, $\overline{v_{\text{in}}} = \overline{v_{\text{out}}}$. In powerful pumps, where $\Delta p > 1MPa$, ΔT may be a measurable quantity, but for the pump configurations investigated in this study, it can be assumed that ΔT is negligible. The magnitude of Δp is a function of the pump's configuration and the surrounding fluid system. Though the pressure head imparted by the surrounding fluid system is typically constant at steady-state, the measured values of p_{in} and p_{out} are time-dependent. The time-dependence is a combined effect of natural pressure fluctuations caused by the operation of the pump, turbulence in the fluid, and noise in the measurement. Illustrative pressure measurements collected at the intake and discharge ports of a centrifugal pump with a two-blade, 12-inch (305 mm) diameter impeller during normal operation are shown in Figure 1.8.

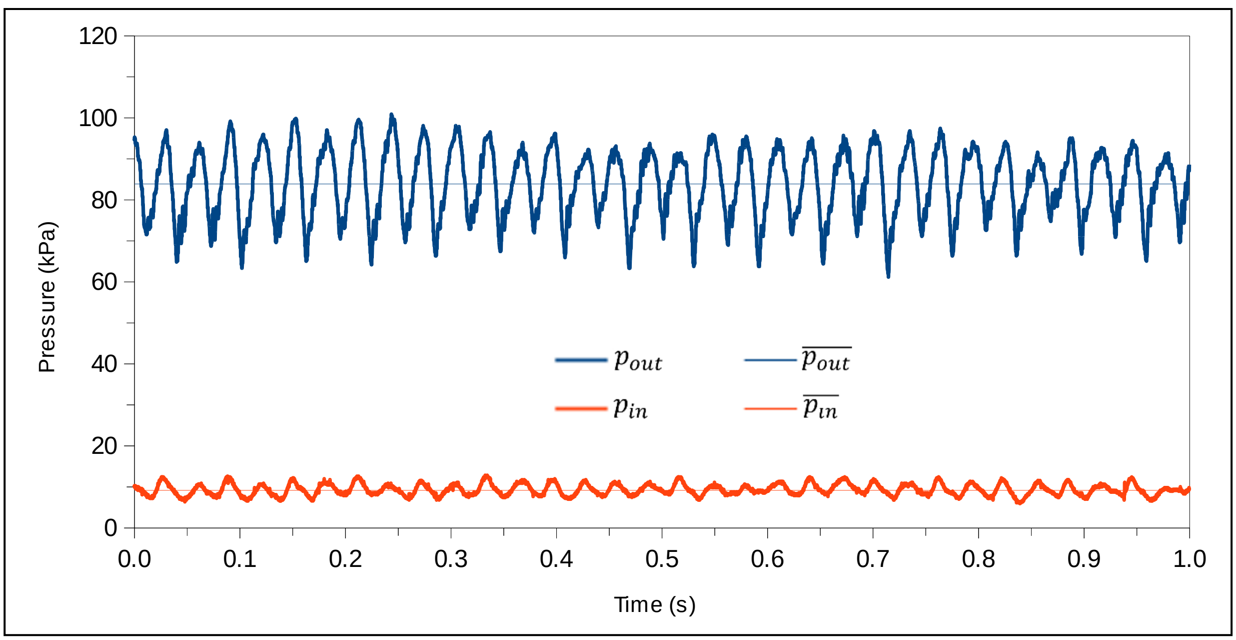


Figure 1.8: Time-varying, experimental pressure measurements at the intake ($\boldsymbol{p_{\text{in}}}$, red line) and discharge ($\boldsymbol{p_{\text{out}}}$, blue line) of a centrifugal pump operating at $\boldsymbol{\omega = 16.4Hz}$. In this example, $\boldsymbol{N_{\text{imp}} = 2}$. The respective sample averages $\boldsymbol{\overline{p_{\text{in}}} = 9.2kPa}$ and $\boldsymbol{\overline{p_{\text{out}}} = 83.9kPa}$ are indicated. The difference $\boldsymbol{\Delta p = 74.7kPa}$. Both measurements exhibit cyclic fluctuations, but the behavior is most apparent in $\boldsymbol{p_{\text{out}}}$.

In practice, when constant pressure measurements are desired – for example, to implement the efficiency determination method in (3), time-averaging would be implemented and the dynamic data discarded. However, the example measurements exhibit a relevant dominant cyclic mode, accompanied by higher frequency fluctuations. It can be reasoned that the fluctuations are more significant in p_{out} because it is measured downstream of the impeller. This behavior parallels results of experimental studies by E. Higham *et al.* [47], Z. Yao *et al.* [60], and Tan *et al.* [61] . Correlating the pressure excursions in p_{out} to the angular position of the impeller θ yields a relative timeline of the pressure events that occur with each rotation of the impeller. This is shown in Figure 1.9.

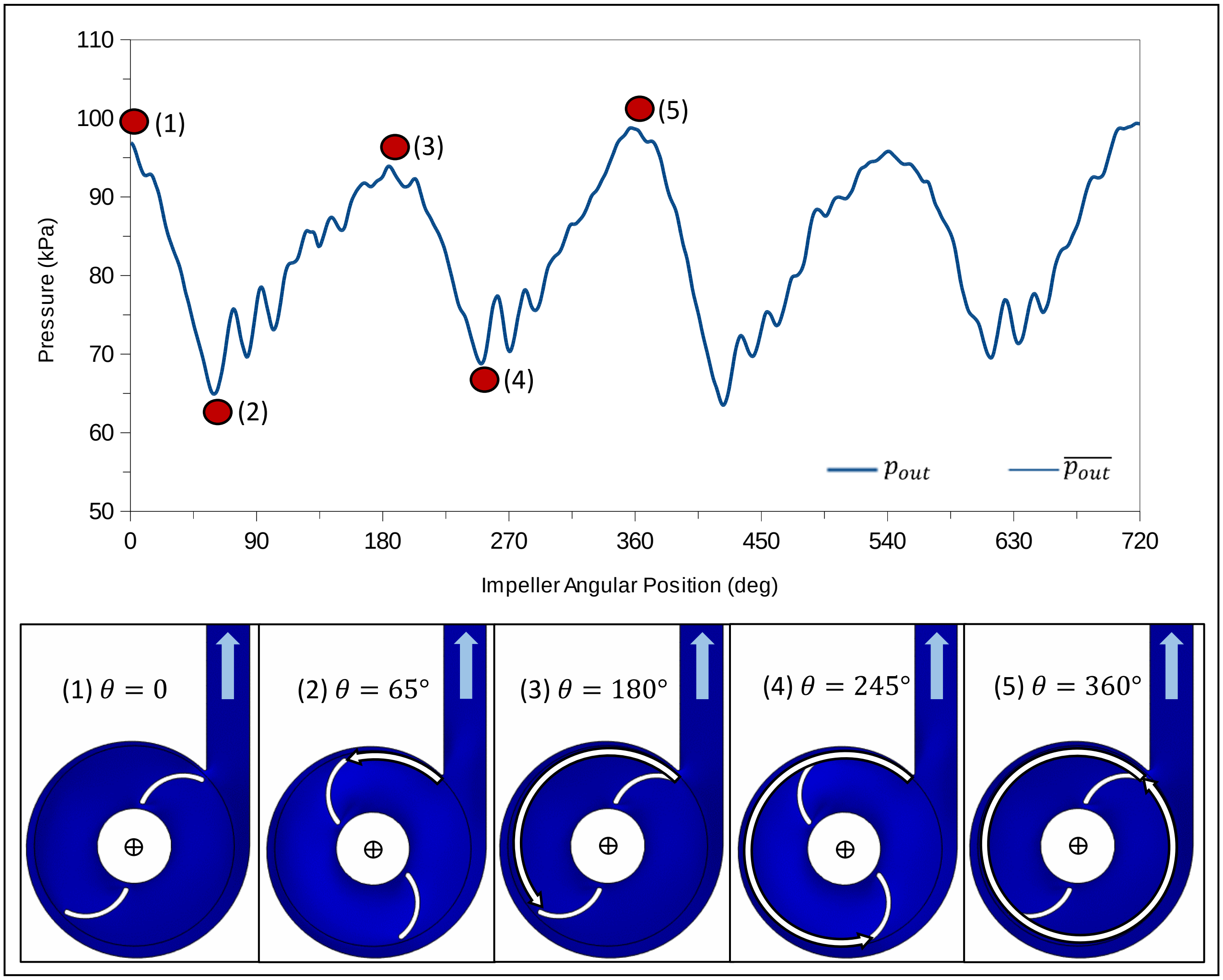


Figure 1.9: $\boldsymbol{p_{\text{out}}}$ as a function of impeller position over two rotations at $\boldsymbol{\omega = 16.4Hz}$. At point (1), the tip of the leading impeller blade has just reached the cutwater. This defines $\boldsymbol{\theta = 0}$. As the impeller rotates to (2), the discharge pressure drops abruptly, only starting to increase again once the second impeller blade starts to

force fluid toward the discharge. The discharge pressure increases until the second impeller reaches the cutwater at (3). A pressure drop follows until the leading impeller blade reaches (4), where the pressure again rises and the cycles continue. Over this time scale, it can be observed that the fluctuations of $\boldsymbol{p}_{\mathbf{out}}$ are not purely sinusoidal, but could be more appropriately generalized as a triangle or sawtooth wave.

Direct interrogation determines that the primary cyclic mode corresponds to the BPF, which is expected. Here, the discharge pressure signal acts as a measure of the fluid-borne vibration. In effect, the pressure transducer functions as a rudimentary hydrophone through which information about the pump's operating state can be deduced, even in the absence of more complex instrumentation.

1.5.2 Gas Entrainment Condition

The gas and liquid constituents of the working fluid in this study are air and water, respectively. The severity of gas entrainment is characterized by the dimensionless void fraction (or equivalently, volume fraction) of air

$$\varphi_{\mathrm{air}} = \frac{q_{\mathrm{air}}}{q_{\mathrm{w}}+q_{\mathrm{air}}}, \tag{9}$$

where q_{air} and q_{w} are the average volumes of air and water passing through the pump over time, respectively. The sum corresponds to the total flow rate $Q = q_{\mathrm{w}} + q_{\mathrm{air}}$.

The distribution of the air phase in the incoming flow in this study can be described as "dispersed bubbly" or "layered bubbly" flow using the morphology conventions described by T. Xie [48]. However, it is assumed that the turbulent mixing contribution of the impeller exceeds the surface tension forces present in larger bubbles considerably, yielding an approximately homogeneous distribution of bubbles with diameters smaller than 5 mm exiting the pump discharge, regardless of the morphology at the intake.

A centrifugal pump's design and surrounding fluid process will ultimately dictate the range of air volume fractions under which it may function, but in a typical application, normal operation coincides with the state of $\varphi_{\mathrm{air}} = 0$ and very severe air entrainment may reach void fractions exceeding 0.05, or 5% [51], [52]. For centrifugal pumps designed to move only liquids, any presence of entrained gas is undesirable. At volume fractions greater than 5%, the performance of a conventional centrifugal pump is likely to have degraded so severely that manual observation would readily perceive the aberration [51]. Thus, higher air volume fractions are not considered.

Adding a significant proportion of entrained air impacts the pressure fluctuations at the pump discharge, as well as the general operation of the centrifugal pump. To illustrate this, the healthy discharge pressure measurement from Figure 1.8 is overlaid with a corresponding measurement at $\varphi_{\text{air}} = 5.0\%$ in Figure 1.10. In the testbed pump, a 5.0% void fraction approaches the threshold of a stall condition. Increasing the volume fraction further creates a loss of pressure and a ceasing of flow.

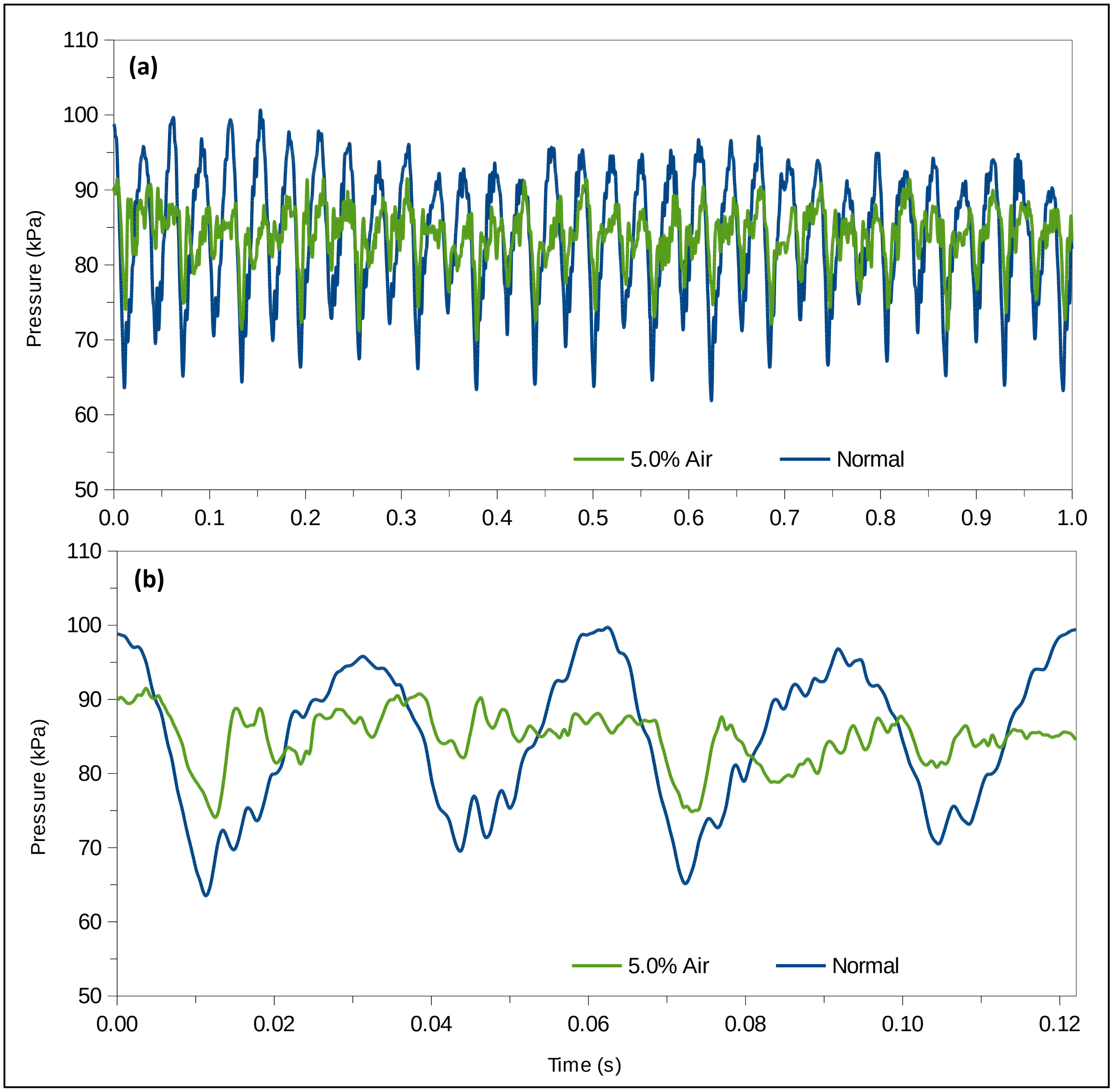


Figure 1.10: a) Experimental 1 s samples of $\boldsymbol{p_{\text{out}}}$ at $\boldsymbol{\omega = 16.4Hz}$ for normal pump operation (blue) and a gas entrainment condition $\boldsymbol{\varphi_{\text{air}} = 5.0\%}$ (green); b) A detail view of the corresponding fluctuations over two impeller rotations. The impeller positions are synchronized for clarity.

Introducing entrained air causes a decrease in the variance between peak and minimum pressures in Figure 1.10a. It can also be observed that the pressure peaks in the sample with entrained air are flattened, which, in turn, indicates that the skewness of the signal has likely changed. Over the shorter time scale in Figure 1.10b, it becomes apparent that the main cyclic mode of pressure sample with air entrainment has deteriorated and high frequency fluctuations have increased. The waveform has also lost much of its triangle-wave-like character in the presence of severe gas entrainment. This preliminary contrast between discharge pressure measurements at $\varphi_{\text{air}} = 0$ and $\varphi_{\text{air}} = 5.0\%$ suggests that more rigorous investigation has the potential to yield substantive trends that can be used to characterize the severity of gas entrainment.

1.5.3 Worn Impeller Condition

There are multiple mechanisms by which an impeller blade can erode. This work focuses on material loss from the impeller tips, radially. Mechanical erosion from the ends of the impeller blades can occur through natural friction wear, abrasion, sudden damage, or intentional manual removal. The intentional case, referred to as impeller trimming, is a modification intended to reduce the fluid volume displaced by a centrifugal pump in the absence of a means to control the rotating frequency. Impeller wear is characterized using the difference

$$\Delta D = D_{\text{o}} - D_{\alpha}\,, \tag{10}$$

where D_{α} is the actual impeller's operating diameter and D_{o} is that of the original, unworn impeller. A schematic of this geometry is shown in Figure 1.11.

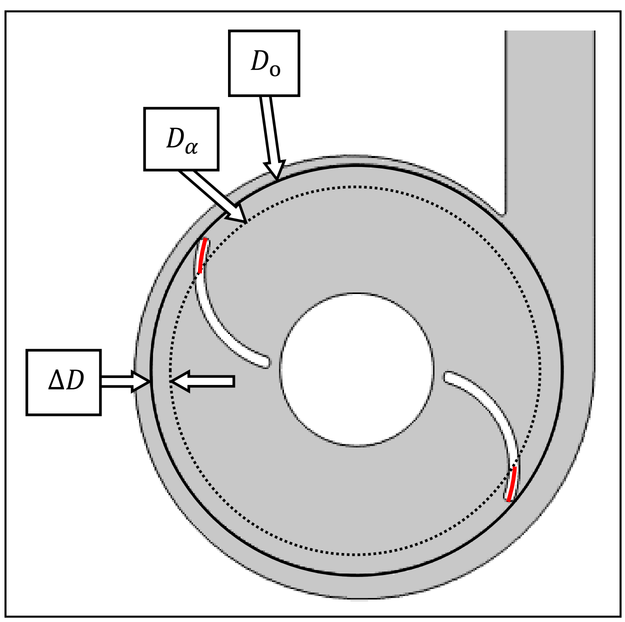


Figure 1.11: Geometry of radial impeller wear (wear region indicated in red). $\Delta\mathbf{D}$ is exaggerated for clarity.

The severity of the tip erosion is defined by impeller loss ratio

$$\alpha_{\text{imp}} = \frac{\Delta D}{D_{\text{o}}}. \tag{11}$$

Substituting (10) into (11) yields the final form,

$$\alpha_{\text{imp}} = \frac{D_{\text{o}} - D_{\alpha}}{D_{\text{o}}}. \tag{12}$$

α_{imp} denotes the relative reduction in impeller diameter from its unworn state. $\alpha_{\text{imp}} = 0$ corresponds to that nominal state, where no material has been eroded from the impeller tips.

The configuration and process of a specific centrifugal pump dictate the range of impeller loss ratios for under which it may continue to operate reasonably normally. In the case of deliberate impeller trimming, impeller loss ratios as high as $\alpha_{\text{imp}} = 0.5$, or 50% may be considered acceptable [5], though the design of the pump would be fundamentally altered, limiting the value of comparing it to its original state. Excluding cases of intentional material removal, which are not the primary focus of this study, it is assumed that realistic impeller loss ratios between 0 and 3% should be considered for characterization, as they are potentially severe enough to cause a deviation in a pumping process, yet may not be readily detectable using more conventional means. Impeller loss ratios exceeding $\alpha_{\text{imp}} = 3\%$ are assumed to be so severe that they would likely be observable through other measurement methods, as demonstrated by M. Matlaka *et al.* [58]. As such, loss ratios beyond 3% are not explicitly investigated in this work.

Generating an illustrative measurement of the dynamic discharge pressure impacts of impeller wear is not trivial, considering it requires permanently damaging a centrifugal pump. Instead, a 2D finite element simulation of a centrifugal pump is created using COMSOL Multiphysics software. The development and validation of this numerical model are discussed in *Chapter 2* of this thesis. Figure 1.12 contrasts simulated discharge pressure measurements for an unworn impeller and the same pump with an impeller loss ratio of $\alpha_{\text{imp}} = 3\%$. The simulation emulates a two-blade, 4-inch (102 mm) diameter impeller rotating at $\omega = 18.75 Hz$.

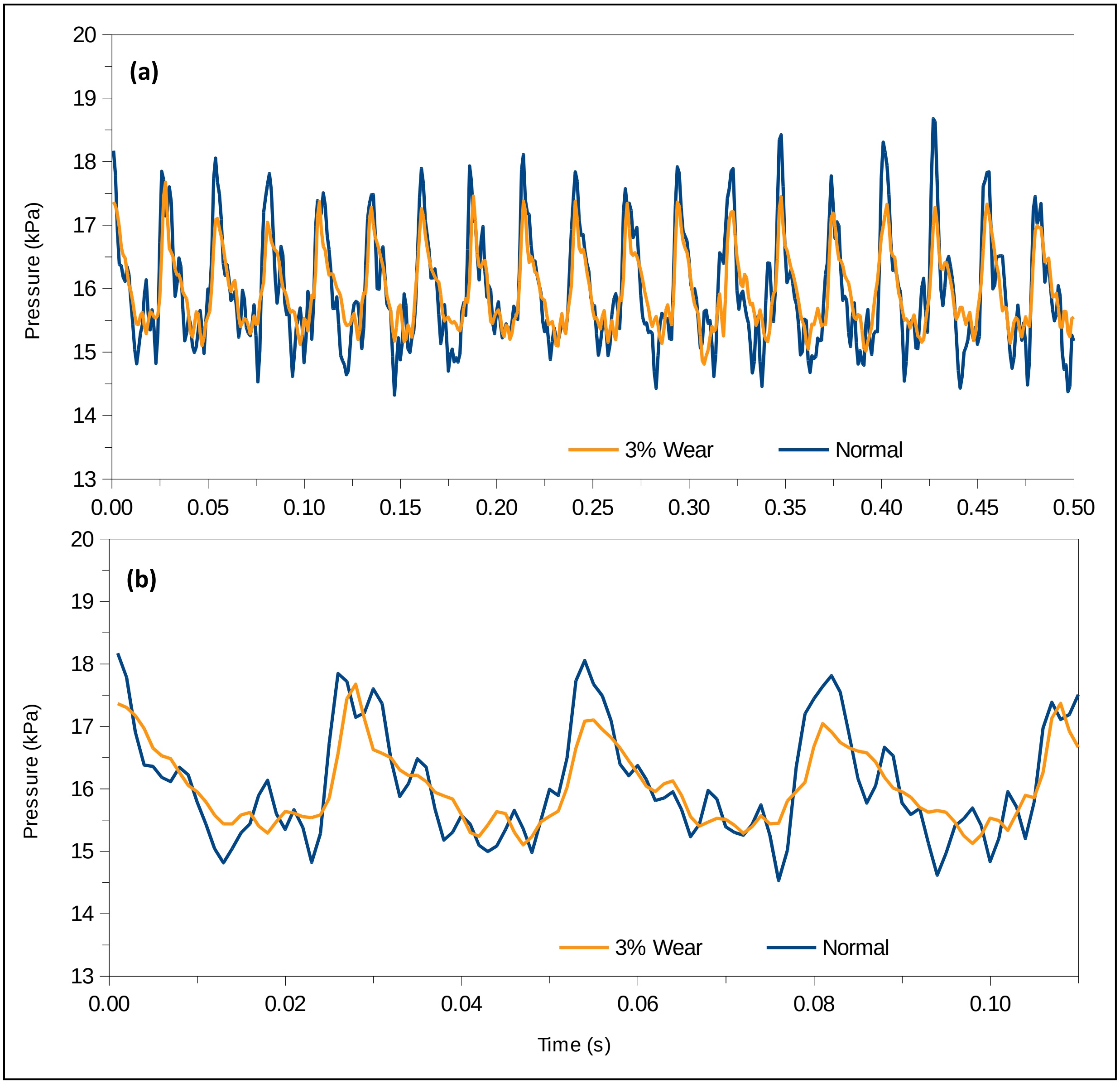


Figure 1.12: a) A simulated 0.5 s sample of $\boldsymbol{p}_{\mathbf{out}}$ for normal pump operation (blue) and an impeller wear condition $\boldsymbol{\alpha}_{\mathbf{imp}} = \mathbf{3.0}\%$ (orange) at $\boldsymbol{\omega} = \mathbf{18.75Hz}$; b) A detail view of the corresponding fluctuations over two impeller rotations. The impeller positions are synchronized for clarity.

The variance between the peak and minimum discharge pressures is reduced in the severe wear condition. The simulated signals show no clear increase in high frequency fluctuations. This is consistent with experimental results reported by A. Suhane [43]. As with the preliminary measurements for air entrainment, the impeller wear simulation exhibits a triangle-wave-like behavior, rather than purely sinusoidal. However, unlike the air entrainment case, the shape of the fluctuations is retained in the presence of radial impeller wear, suggesting that the changing shape of the waveform could be a telltale feature to differentiate between air entrainment or impeller wear.

1.6 Objectives

The research discussed in the previous sections suggests a pressing need for affordable, accurate, and readily implementable methods to monitor gas entrainment and impeller wear in centrifugal pumps. Measurement of fluid pressure fluctuations at the pump discharge, in conjunction with machine learning methods, offers potential in this respect. Considering this, the objectives of this research are as follows:

(1) Statistical Features: Propose a set of statistical measures to quantify the severity of the target phenomena using dynamic discharge pressure.

(2) Correlation: Assess the relevance of these measures as related to detecting the target phenomena.

(3) State Prediction: Develop a classification technique to detect with at least 90% accuracy the occurrence of a) air entrainment exceeding a 2% void fraction and b) radial wear exceeding 1.5% of the impeller diameter in an operating centrifugal pump.

(4) Refined Characterization: Demonstrate refined characterization of the target phenomena using multiple severity classes and continuous-value regression.

(5) Accessibility: Implement the detection method using a single pressure transducer that costs less than $200 and has a rise time no faster than 2 ms.

(6) Implementation: Suggest protocols for industrial implementation of the method.

The following chapters describe the design, characterization, and evaluation of the system. *Chapter 2* defines the numerical modeling, experimental materials, and methods employed to generate discharge pressure measurements of the target phenomena. *Chapter 3* contains the analytical background for the statistical measures, analysis method, and classification models. *Chapter 4* contains the experimental results and corresponding analysis. Finally, *Chapter 5* draws conclusions and describes potential further applications of this work.

2. MATERIALS AND EXPERIMENTAL METHODS

The studies in this work are conducted in two stages. First, we employ numerical modeling to generate dynamic pressure measurements corresponding to normal, intermediate, and extreme states for each of the target phenomena. Contrasting these disparate conditions provides a basis from which to derive relevant measures to quantify the target conditions and propose classification algorithms. The workflow of these initial studies is shown in Figure 2.1 and discussed in *Chapter 3*.

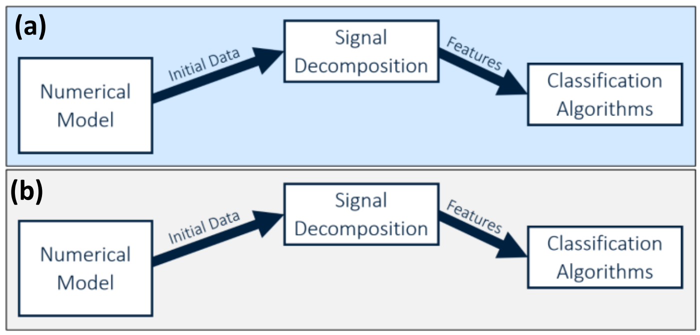


Figure 2.1: Workflow of the preliminary studies for a) gas entrainment and b) impeller wear. The initial simulated data is used to evaluate signal decomposition techniques and propose machine learning methods.

In the second stage (Figure 2.2), we generate comprehensive sets of dynamic pressure measurements for each target condition. Air entrainment data is generated using experimental methods. Impeller wear pressure data is simulated using the numerical model. These complete sets of dynamic pressure measurements are then used to demonstrate and validate the classification methods in *Chapter 4*.

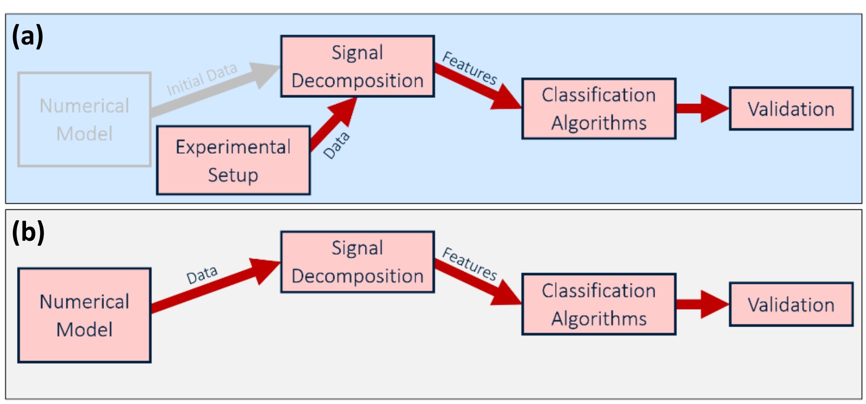


Figure 2.2: Workflow of the primary studies for a) gas entrainment and b) impeller wear. Experimental data is used to validate gas entrainment classification, whereas numerical data is used for impeller wear.

2.1 Simulation

A numerical model of a rotating centrifugal pump is developed using COMSOL Multiphysics software (v. 5.5, b. 359, COMSOL Inc. 2019) to assess the pressure dynamics associated with gas entrainment and impeller wear. The model applies finite element analysis (FEA) to perform time-dependent simulations of each of the target conditions, across a range of operating states, to generate an array of characteristic dynamic pressure signals at the pump discharge.

2.1.1 Model Configuration

The simulated pump is based on an Ampco AC114 centrifugal pump, shown in Figure 1.1. To reduce the computational complexity to a practical scale, the modeled geometry requires a series of simplifications. First, being that centrifugal pumps contribute kinetic energy to the working fluid in an inherently radial manner, in the plane of the impeller, it is assumed that the dominant modes of pressure fluctuation and turbulence will also be in-plane. The system is therefore modeled in 2D. Secondly, to minimize the required quantity of finite elements, detailed structures that are not strictly essential to pump operation are removed or simplified. This includes a) reducing the number of impeller blades from three to two, b) rounding sharp corners at the impeller ends and the cutwater, c) generalizing the impeller blade contours as radial, rather than a more complex spline, and d) minimizing the length of the discharge channel, to the extent that its boundary condition does not interfere with the pressure fluctuations. The geometry of the simulated pump is shown in Figure 2.3.

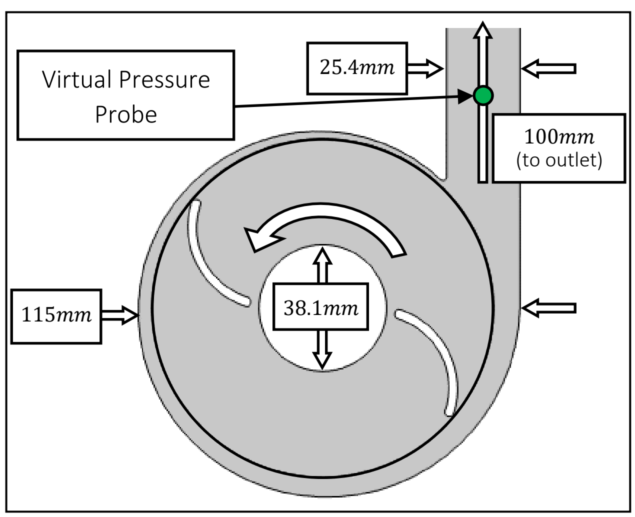


Figure 2.3: Geometry of the simulated centrifugal pump. The virtual pressure probe (green dot) is positioned in the center of the discharge, 50 mm from the pump outlet.

In the nominal model configuration, $D_\mathrm{o} = 10.2mm$. The minimum clearance between the impeller and cutwater is 0.75 mm.

The operating area of the simulation is composed from two domains. First is the dynamic domain, which comprises the fluid inlet and the area of the rotating impeller blades. Second is the static domain, which encompasses the volute wall and discharge channel. The rotational frequency ω is applied as the input to the dynamic domain, therefore the position of the impeller throughout the simulation is time-dependent. The circular intersection of the outer surface of the rotating domain and inner surface of the static domain is designated as an "identity pair," meaning the solver treats the junction as a coincident surface across which continuity and fluid transport are to be maintained, even as the impeller domain rotates.

Flow through the simulated pump is bounded by the conditions at the pump intake and discharge, which, in a physical system, are a function of ω. At the discharge, a fully developed flow condition with a specified average pressure is applied. The average pressure increase Δp is chosen to be a fixed hydraulic head of 1.5 m (14.7 kPa).

The flow rate Q is estimated using the manufacturer's data sheet for the Ampco AC114 centrifugal pump [62]. Flow rates of 50 l/min at the maximum impeller speed of $\omega_{\mathrm{max}} = 28.75Hz$ and 15 l/min at the minimum impeller speed $\omega_{\mathrm{min}} = 8.75Hz$ are assumed, with linear behavior between. The flow rates are considered only as a representative approximation, given that the configuration of the impeller blades has been simplified for modeling. The average outlet velocity $\overline{\mathrm{v}_{\mathrm{out}}}$ is determined by dividing Q by the outlet area.

At the pump inlet, a normal inflow velocity $\overline{\mathrm{v}_{\mathrm{in}}}$ at atmospheric pressure is specified. The magnitude of $\overline{\mathrm{v}_{\mathrm{in}}}$ is determined by multiplying $\overline{\mathrm{v}_{\mathrm{out}}}$ by the ratio between the diameter of the pump outlet and the circumference of the model's inlet (rather than their respective areas, because in 2D, the outlet flow is in plane and the inlet is not). With $\overline{p_{\mathrm{in}}} = 0kPa$ (gauge), $\Delta p = \overline{p_{\mathrm{out}}}$.

Samples of p_{out} are collected at a virtual probe point 50 mm below the model's outlet boundary, as indicated in Figure 2.3. The probe is positioned to avoid the turbulent effects in the immediate area of the rotating impeller, but sufficiently far from the fluid exit boundary condition to avoid damping effects from the boundary condition. The positional impact of the probe placement is shown in Appendix B1. To attain quasi-steady state operation, ω, Δp, and Q are linearly ramped up from rest over a buffer period of 1.5 s. Simulated pressure data from this transient period is not considered.

Across all the simulations, dynamic pressure samples are collected over a total time duration $T_s = 0.5s$ at a sampling frequency $f_s = 10^3 Hz$. This yields $N = 500$ pressure data points within each measurement. The sampling interval $t_s = 1/f_s = 10^{-3} s$. The duration of T_s is selected to provide frequency resolution $\Delta f = 1/T_s = 2Hz$, down to the minimum frequency of interest $f_{\min} = 2Hz$.

The dynamics of a physical pressure transducer are not reflected in the virtual probe measurement. To conservatively avoid the potential influence of any fluctuations exceeding what would be experimentally observable using a pressure sensor with a rise time of 2 ms, frequencies exceeding 400 Hz are filtered from the simulated signal with a low-pass filter (LPF). 400 Hz is the maximum frequency of interest $f_{\max}$. Any potential fluctuations resulting from coupling changes – that is, whether the transducer diaphragm is in contact with the liquid, gas, or both phases at a given instant – cannot be replicated using the numerical model.

2.1.2 Numerical Methods

The centrifugal pump simulation incorporates two physics models; a flow modeling method and a multi-phase method.

Turbulence Model

Flow through a centrifugal pump is assumed to have regions with a high Reynolds number, so a numerical model that considers turbulence is employed. The Realizable k-ε turbulence model has been applied successfully in related studies [63] and demonstrated to be particularly suited for applications with rotating flows [64]. It is the turbulence model selected for this research.

The Realizable k-ε approach solves transport equations for the turbulent kinetic energy of the liquid phase k_l and the turbulent energy dissipation rate of the liquid phase ε_l using the realizability conditions

$$\overline{u_i^2} \geq 0 \tag{13}$$

and

$$\frac{\overline{u_i u_j}^2}{\overline{u_i^2}\,\overline{u_j^2}} \leq 1 \tag{14}$$

where u_i and u_j are the unfiltered components of the time-averaged flow velocity vector $\boldsymbol{u}$. The transport equation for k_l,

$$\rho \frac{\partial k_l}{\partial t} + \rho(\boldsymbol{u} \cdot \nabla) k_l = \nabla \cdot \left[\left(\mu + \frac{\mu_{Tl}}{\sigma_k} \right) \nabla k_l \right] + P_k - \rho \varepsilon_l \tag{15}$$

is equivalent to that of the standard k-ε method [64]. In this steady-state application, the time rate of change of the turbulent kinetic energy of the liquid phase $\frac{\partial k_l}{\partial t} = 0$. The dissipation rate ε_l is calculated by the transport equation

$$\rho \frac{\partial \varepsilon_l}{\partial t} + \rho(\boldsymbol{u} \cdot \nabla) \varepsilon_l = \nabla \cdot \left[\left(\mu + \frac{\mu_{Tl}}{\sigma_\epsilon} \right) \nabla \varepsilon_l \right] + C_{\varepsilon 1} \rho S \varepsilon_l - C_{\varepsilon 2} \rho \frac{\varepsilon^2}{k_l + \sqrt{\nu_l \varepsilon_l}}, \tag{16}$$

where the time rate of change of the turbulent dissipation rate of the liquid phase $\frac{\partial \varepsilon_l}{\partial t} = 0$.

In (15) and (16), the total fluid density ρ and dynamic viscosity μ are determined by the two-phase model (see (21) and (22)). Kinematic viscosity ν_l corresponds to the liquid fluid phase (water). P_k represents the turbulent kinetic energy production term. S is the strain-rate tensor. The adjustable constants $\sigma_k = 1.0$, $\sigma_\epsilon = 1.2$, and $C_{\varepsilon 2} = 1.9$ are used [63]. The turbulent eddy viscosity of the liquid phase

$$\mu_{Tl} = C_\mu \rho_l \frac{{k_l}^2}{\varepsilon_l} \tag{17}$$

contains a scalar term C_μ that is not constant. The computations for C_μ, P_k, S, the turbulent production rate scalar $C_{\varepsilon 1}$, and their associated calculations are shown in Appendix A1.

Continuity is determined using the Reynold's Averaged Navier-Stokes (RANS) turbulence equations

$$\rho \frac{\partial \boldsymbol{u}}{\partial t} + \rho(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = \nabla \cdot \{-p\boldsymbol{I} + \mu[\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T]\} + F - \rho \boldsymbol{g} \tag{18}$$

and

$$\nabla \cdot \rho \boldsymbol{u} = 0\,, \tag{19}$$

where the fluid is approximated to be incompressible and Newtonian. $\boldsymbol{I}$ is the identity tensor, F is the volume shear force, and $\boldsymbol{g}$ is the gravitational field. The wall boundaries of the simulated pump, including the rotating surfaces of the impeller blades, are assumed to be non-slip.

Multi-phase Model

In the multi-phase portion of the numerical model, it is assumed that the dynamic pressure contribution of air bubble coalescing and breakup is non-trivial and must be accounted for, despite the non-trivial computational cost of calculating individual phase boundaries in a time-dependent study. Simpler two-phase mixing models, such as the Euler-Euler method used by J. Zhang *et al.* [8] and the distributed bubbly flow technique applied by K. Minemura and T. Uchiyama [65] are not considered suitable. The Euler-Euler method assumes that the fluid phases are interpenetrating and calculates pressures using mixture averaging, which is unlikely to generate satisfactory resolution of pressure fluctuations around the gas phase. The bubbly flow technique by Minemura requires that the volume of entrained gas be small and homogeneously distributed enough for coalescing and fragmentation to be neglected. Dynamic pressure considerations notwithstanding, it is not clear that this is a justifiable assumption in flows with φ_{air} exceeding a few percent.

To achieve detailed spatial and temporal resolution of the air-water interface while keeping computational cost manageable, a level set two-phase flow model is employed [66]. The level set method is intended for simulating the boundary between two immiscible phases (e.g. liquid water and air). The technique defines a level-set variable $0 \leq \phi \leq 1$, where $\phi = 0$ and $\phi = 1$ correspond to the pure water and pure air domains, respectively, with a transition region of thickness ϵ between them. Here, a quadratic transition is employed and ϵ is defined as one half the length of the largest mesh element in the simulation. The fluid interface is defined as the iso-contour $\phi = 0.5$. The motion of the interface is calculated by solving

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla\phi = \zeta\nabla \cdot \left[\epsilon\nabla\phi - \phi(1-\phi)\frac{\nabla\phi}{|\nabla\phi|}\right], \tag{20}$$

where ζ is a dimensionless tuning parameter to control the numerical stability of the function. The total fluid density

$$\rho = \rho_{\mathrm{water}} + (\rho_{\mathrm{air}} - \rho_{\mathrm{water}})\phi \tag{21}$$

and the total dynamic viscosity

$$\mu = \mu_{\mathrm{water}} + (\mu_{\mathrm{air}} - \mu_{\mathrm{water}})\phi \tag{22}$$

apply to transport equations (15) and (16). The void fraction of air entering the system is specified by the inlet condition. Surface tension is considered as well.

2.1.3 Mesh

The simulation mesh is optimized to balance dynamic accuracy and computational time. Element size is controlled using COMSOL Multiphysics' integrated meshing function, which increases the density in areas with complex geometry.

Mesh independence is validated by simulating the condition $\varphi_{\mathrm{air}} = 5.0\%$ and $\omega = 18.75Hz$ with successively increasing element density, until minimal influence is observed. The influence of mesh density becomes negligible at 10^4 elements. However, at this density, generating the corresponding 0.5 s pressure sample takes approximately seven hours. For a single simulation, this is not insurmountable, but for building extensive sets of simulated conditions, it is problematic. Reducing the mesh density to 10^3 elements shortens the corresponding computation time by a factor of five, while retaining the prominent fluctuation characteristics (as validated in Section 2.1.4). Thus, this density is selected for the model.

2.1.4 Validation

The numerical model is validated for ordinary operation and gas entrainment conditions against the preliminary experimental measurements from Figure 1.10. Overlaying the simulated data onto the experimental samples yields the correlations in Figure 2.4.

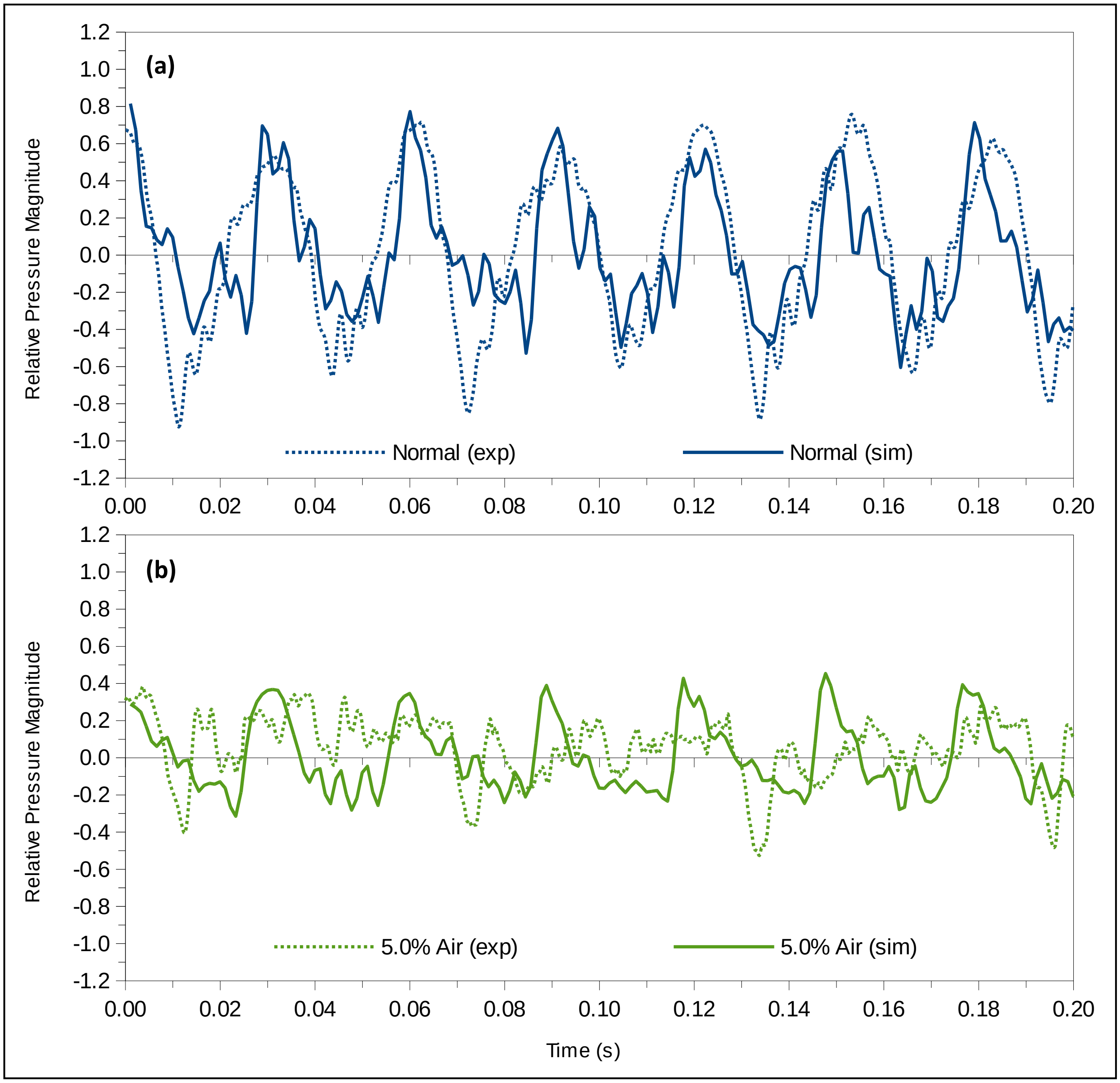


Figure 2.4: a) Contrasting experimental (dotted) and simulated (solid) measurements of $\boldsymbol{p}_{\mathbf{out}}$ for normal pump operation at $\boldsymbol{\omega} = \mathbf{16.4}\boldsymbol{Hz}$. Each sample is 0.5 s total and normalized to its maximum value. The first 0.2 s are shown for clearer interpretation of the waveforms; b) Contrasting experimental (dotted) and simulated (solid) measurements of $\boldsymbol{p}_{\mathbf{out}}$ for the gas entrainment condition $\boldsymbol{\varphi}_{\mathbf{air}} = \mathbf{5.0}\%$. Each gas entrainment condition is normalized to the maximum value of its corresponding healthy condition from Figure 2.4a to show the relative reduction in variance.

The experimental and simulated discharge pressure signals in Figure 2.4a exhibit similarities. Both waveforms are approximately triangular with a dominant fluctuation mode at the BPF and contribution from higher frequencies. Calculating their Pearson correlation coefficient yields an agreement of 0.83

(see (A10) in Appendix A2 for calculation). In both the experimental and simulated signals, the inclusion of a 5.0% void fraction of air causes a decrease in the variance of the pressure fluctuations and disruption of the main fluctuation modes (Figure 2.4b). Together, these advocate that the numerical model provides a satisfactory representation of the primary characteristic elements of a physical centrifugal pump.

2.2 Experimental Apparatus

Dynamic discharge pressure data from a physical centrifugal pump is used to validate the simulation dynamics and provide reference data by which to evaluate classification methods. Physical pressure data is collected for normal and air entrainment conditions only.

2.2.1 Pump and Fluid System

Experimental measurements are collected on the centrifugal pump powering the pilot-scale pulp refining loop (Pilot Plant) in the Paper and Pulp Centre (PPC) at The University of British Columbia (UBC), shown in Figure 2.5.

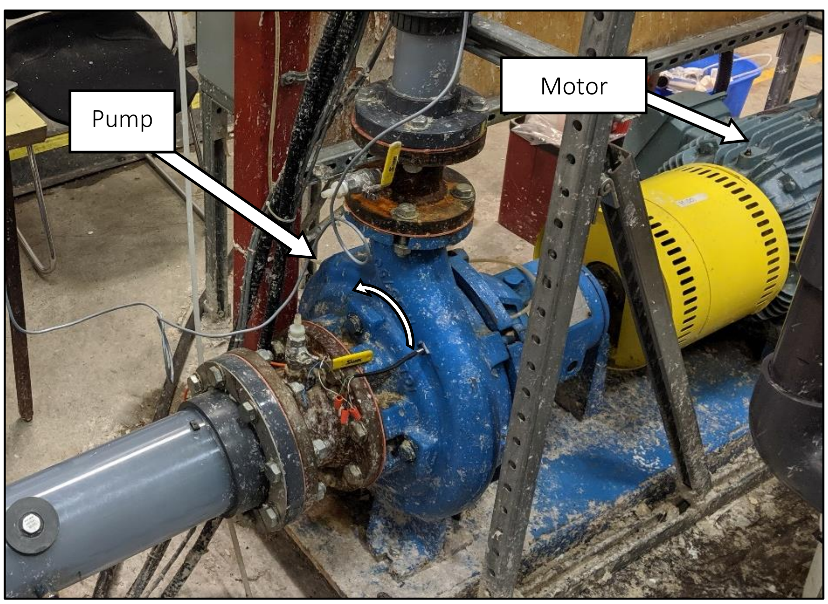


Figure 2.5: The 30 kW centrifugal pump used as a testbed throughout the experimental studies.

The testbed pump is a 40 HP (30 kW) model manufactured by Westcan Industries Ltd. The pump contains a 2-blade open impeller with a diameter of 12 in (305 mm). Driving power is delivered by a close-coupled three-phase electric motor, connected to a variable-frequency drive (VFD) with integrated power and rotating frequency sensing. The pump is capable of displacing nearly 6,000 liters per minute at 3250 RPM, though the construction of the surrounding fluid system limits safe operation to approximately 1,400 liters per minute, which corresponds to 1,125 RPM ($\omega = 18.75Hz$) under minimal pressure load. The minimum impeller speed required to propel fluid through the system, from stationary, is approximately 525 RPM ($\omega = 8.75Hz$).

A schematic of the adjoining pump loop is shown in Figure 2.6.

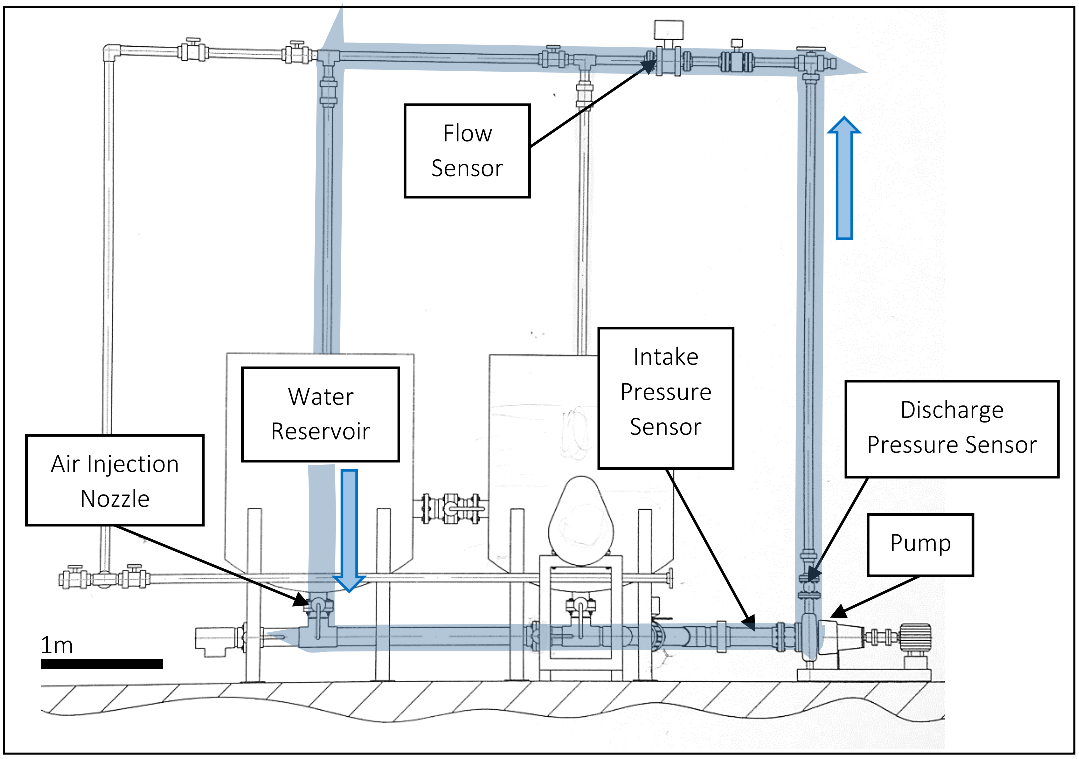


Figure 2.6: A schematic of the fluid loop in the pilot-scale pulp refining plant at UBC.

The loop comprises a 5,000 liter water reservoir, leading to a 4 m horizontal segment of pipe which feeds the pump intake. The pump discharge feeds into a 5 m section of vertical pipe, which then traverses back and down to the water tank. The water re-enters the tank below the surface, creating a closed fluid loop.

Air is injected into the fluid flow at the base of the reservoir using an extended compressed air nozzle with a perforated tip attachment. The reservoir is necessarily large to allow time for the entrained air to dispel to the atmosphere prior to recirculation, preventing accumulation over time. The air flow rate is measured using a rotameter and controlled by a manual gate valve.

2.2.3 Sensors

As discussed in the opening sections of this thesis, it is required that pressure measurements be collected using a transducer that would be affordable and readily attainable in industrial practice. While there exist higher-performance sensors that would potentially provide more refined pressure measurements, the intent is to demonstrate characterization of the target phenomena using a sensor available to any centrifugal pump operator. The objectives in Section 1.6 set this threshold to be under \$200, with a rise time no faster than 2 ms.

A Honeywell MLH series piezoresistive pressure transducer (model MLH050PGB06A, Figure 2.7) [44] is employed to measure p_{out}. The sensor is mounted through the wall of the discharge pipe, 0.5 m from the pump outlet. The transducer has an operating pressure range up to 345 kPa and an accuracy of ±0.25% of full-scale. Full-scale output is 4.5 V and the excitation voltage is 5 V. The response time is specified by the manufacturer as 2 ms. The cost per transducer is approximately \$115 USD.

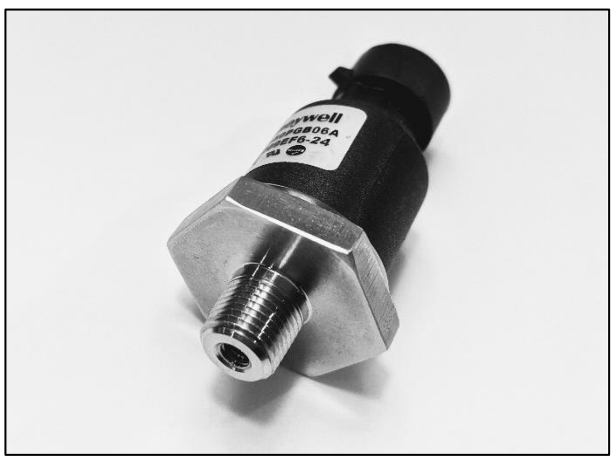


Figure 2.7: A Honeywell MLH050PGB06A pressure transducer.

A series of secondary sensors are used to control the apparatus and provide supplementary data for the analysis. The intake pressure p_{in} is measured by an additional Honeywell MLH050PGB06A pressure transducer at the pump intake. Flow rate Q is determined using a Rosemount 8700 Series magnetic flowmeter [67]. The values of W_{in} and ω are determined using the wattmeter integrated into the pump's variable frequency drive (VFD), which is a Baldor VS1PF-NM1E. q_{air} is measured using an Omega FL-203 rotameter. The operating parameters and associated accuracies for each sensor are shown in Table 2.1.

Table 2.1: Sensor Parameters

Measurand	Symbol	Manufacturer	Model No.	Type	Range	Accuracy
Discharge Pressure	p_{out}	Honeywell	MLH050PGB06A	Piezoresistive	0-350 kPa (gauge)	±0.25% at f.s.
Intake Pressure	p_{in}	Honeywell	MLH050PGB06A	Piezoresistive	0-350 kPa (gauge)	±0.25% at f.s.
Rotating Frequency	ω	Baldor	VS1PF-NM1E	Integrated	0-3500 Hz	Unspecified
Input Power	W_{in}	Baldor	VS1PF-NM1E	Integrated	0-30 kW	Unspecified
Flow Rate	Q	Rosemount	8707	Magnetic	3500 l/min	±0.5% at f.s.
Air Flow Rate	q_{air}	Omega	FL-203	Rotameter	11.3-140 l/min	±5% at f.s.

2.2.4 Data Collection

Experimental measurements of p_{out} are collected using a Measurement Computing USB-1208HS data acquisition (DAQ) device [68]. The DAQ is controlled by a custom program, written in Python 3.5, utilizing functions from the Measurement Computing Python library. In each measurement, the data collection interval $T_{\text{s}} = 1s$, allowing for scrutiny of frequencies as low as 1 Hz, which is below f_{min}. This yields $\Delta f = 1Hz$. The sampling frequency $f_{\text{s}} = 10^4 Hz$ and $t_{\text{s}} = 10^{-4} s$. A total of $N = 10^4$ data points are generated within each experimental measurement. To exclude any potentially dubious dynamic behavior attained near the transducer's response limit of 500 Hz, each measurement is filtered by a LPF with a cutoff at 400 Hz prior to analysis. Oversampling is employed to resolve the wave shape and peak values of the pressure fluctuations.

The static measurements of Q, W_{in}, ω, and q_{air} are manually recorded for each sample. Where employed, the dynamic measurements of p_{in} are collected in the same manner as p_{out}.

2.3 Gas-entrainment Study

The goals of the gas entrainment study are to a) evaluate the associated pressure phenomena using simulation, b) generate the phenomenon experimentally to validate the analysis from the simulated phase and provide training/testing data for classification, and c) apply the classification algorithms to characterize the severity of gas entrainment in the physical system. The primary objective is to differentiate experimental operating states into "acceptable" or "severe" conditions (i.e. binary classification) using the condition $\varphi_{\text{air}} > 2\%$ as the criterion for "severe." For this basic characterization, the threshold for a satisfactory classification success rate is 90%. The second objective is to categorize the operating states into multiple classes, corresponding to $0 \leq \varphi_{\mathrm{air}} < 1\%$, $1\% \leq \varphi_{\mathrm{air}} < 2\%$, etc. and appraise the accuracy. The final objective is to implement a prediction model that yields a continuous-value prediction of φ_{air} and evaluate the prediction error.

The study is performed in two portions. First, the numerical model is used to generate preliminary measurements on which an analysis method and classification approach can be established. Following that, a comprehensive set of experimental state measurements is collected, allowing for validation and refinement of the analysis and classification methods.

2.3.1 Preliminary Simulation

The ranges of air volume fractions investigated by Schäfer ($\varphi_{\mathrm{air}} = 0 - 5\%$) [51] serve as an initial reference for defining a realistic range of gas entrainment conditions. It is assumed that void fractions of 0, 2.5%, and 5.0% reflect normal, moderate, and severe gas entrainment, respectively. Referencing the range of safe impeller speeds from the physical testbed pump ($8.75Hz \leq \omega \leq 18.75Hz$), simulations are conducted at $\omega = 8.75Hz$, $13.75Hz$, and $18.75Hz$ for healthy, moderate, and severe states, yielding a total of nine initial reference states.

2.3.2 Classification States

An extensive map of experimental operating conditions is necessary to generate a data set suitable for the application of machine learning. To achieve this, the operable[4] range of ω values for the testbed pump is divided into 0.5 Hz increments, from $\omega = 8.9 - 16.9Hz$[5] (17 impeller speeds in total). Airflow into the system is divided into 24 steps at fixed flow rates, from $q_{\mathrm{air}} = 0 - 68.0l/min$. This yields 408 total experimental air entrainment states.

[4] A small safety margin is added at the high end of the operable values of ω. This is done to conservatively avoid any potentially damaging behavior that might result from the addition of gas bubbles to the fluid.

[5] Arbitrary non-integer values of ω are used to minimize the likelihood of overlap with 60 Hz electrical noise or any of its associated harmonics during analysis.

2.4 Impeller Wear Study

The impeller wear study is conducted using a method similar to the gas entrainment study. The intent is to generate a small initial set of simulated wear conditions, evaluate the relevant behaviors, and then generate a comprehensive set of dynamic pressure measurements, corresponding to a variety of impeller wear conditions, on which to apply machine learning. However, to circumvent the need to damage the physical centrifugal pump to produce this comprehensive set, the pressure samples are instead generated using the numerical model.

The primary objective is to distinguish radial impeller erosion into "acceptable" and "severe" states using the threshold $\alpha_{\mathrm{imp}} > 1.5\%$ to delineate "severe." The performance target for this binary classification is 90%. The second objective is to differentiate the conditions into multiple classes, corresponding to $0 \leq \alpha_{\mathrm{imp}} < 1\%$, $1\% \leq \alpha_{\mathrm{imp}} < 2\%$, etc. As with the air entrainment study, the final objective is to demonstrate a model that yields a continuous-value prediction of α_{imp}, and assess its prediction accuracy.

The impeller wear study is performed in two segments. The numerical model is initially used to generate a small set of pressure measurements to permit development of an analytical basis for classifying radial impeller erosion. After, an extensive set of simulated conditions is generated to be used for training, testing, and validation of the classification technique.

2.4.1 Preliminary Simulation

The work by A. Suhane suggests that the adverse impacts of radial impeller clearance become conspicuous when the nominal gap between the impeller and cutwater grows by a factor of five [43]. Applying this factor to the nominal clearance in the simulated pump ($D_{\mathrm{o}} = 0.75mm$) yields a radial loss of approximately 3 mm. From (10), the corresponding impeller loss ratio can be calculated as $\alpha_{\mathrm{imp}} = 3\%$. This is used as the initial reference for severe radial impeller erosion. Impeller loss ratios for normal, moderate, and severe radial erosion are defined as 0, 1.5%, and 3.0%, respectively. Evaluating each of these wear conditions at $\omega = 8.75Hz$, $13.75Hz$, and $18.75Hz$ yields nine initial reference states.

2.4.2 Classification States

Samples from a wide range of states are necessary to apply machine learning to classify radial impeller wear. In this simulated case, the range of impeller speeds are not bounded by the applied safety constraints of the physical system, allowing for a wider range of ω values to be investigated. Incrementing in 5 Hz steps, a collection of simulated measurements is compiled for $8.75Hz \leq \omega \leq 28.75Hz$ (five speeds total). The impeller loss ratio is incremented in 0.25% steps, from $0 \leq \alpha_{\text{imp}} \leq 3.0\%$ (thirteen wear states total). This yields a total of 65 simulated states.

2.5 Combined Study

A preliminary investigation is conducted to examine the feasibility of distinguishing gas entrainment, impeller wear, or the simultaneous presence of both using dynamic pressure measurements. The following simulated states are considered at $\omega = 8.75Hz$, $13.75Hz$, and $18.75Hz$, a) normal operation, b) severe gas entrainment ($\varphi_{\text{air}} = 5.0\%$), severe radial erosion ($\alpha_{\text{imp}} = 3.0\%$), and the simultaneous presence of both conditions in severe form. The objective is to provide initial context on whether the proposed diagnostic method is potentially suitable for classifying such a condition.

3. PRINCIPLES AND DESIGN

This chapter comprises the analytical theory and design of the classification method. Here we discuss the techniques used to reduce dynamic discharge pressure measurements into more comprehensible sets of statistical measures. Justification for these measures is provided using data from the simulations described in Sections 2.3.1 and 2.4.1. Following that, three classification architectures are discussed, each with increasing refinement in how precisely it predicts the severity of the corresponding phenomenon.

3.1 Reference Signals

An array of discharge pressure measurements $\boldsymbol{P}_{\mathbf{out}}$ is simulated for each of the conditions in Sections 2.3.1 and 2.4.1. Each $\boldsymbol{P}_{\mathbf{out}}$ comprises a list of N discrete measurements of $p_{\mathbf{out}}$ at the given operating state, each having a pressure magnitude x_n at index n, where $1 \leq n \leq N$. A portion of the reference simulations for gas entrainment and impeller wear is shown in Figure 3.1.

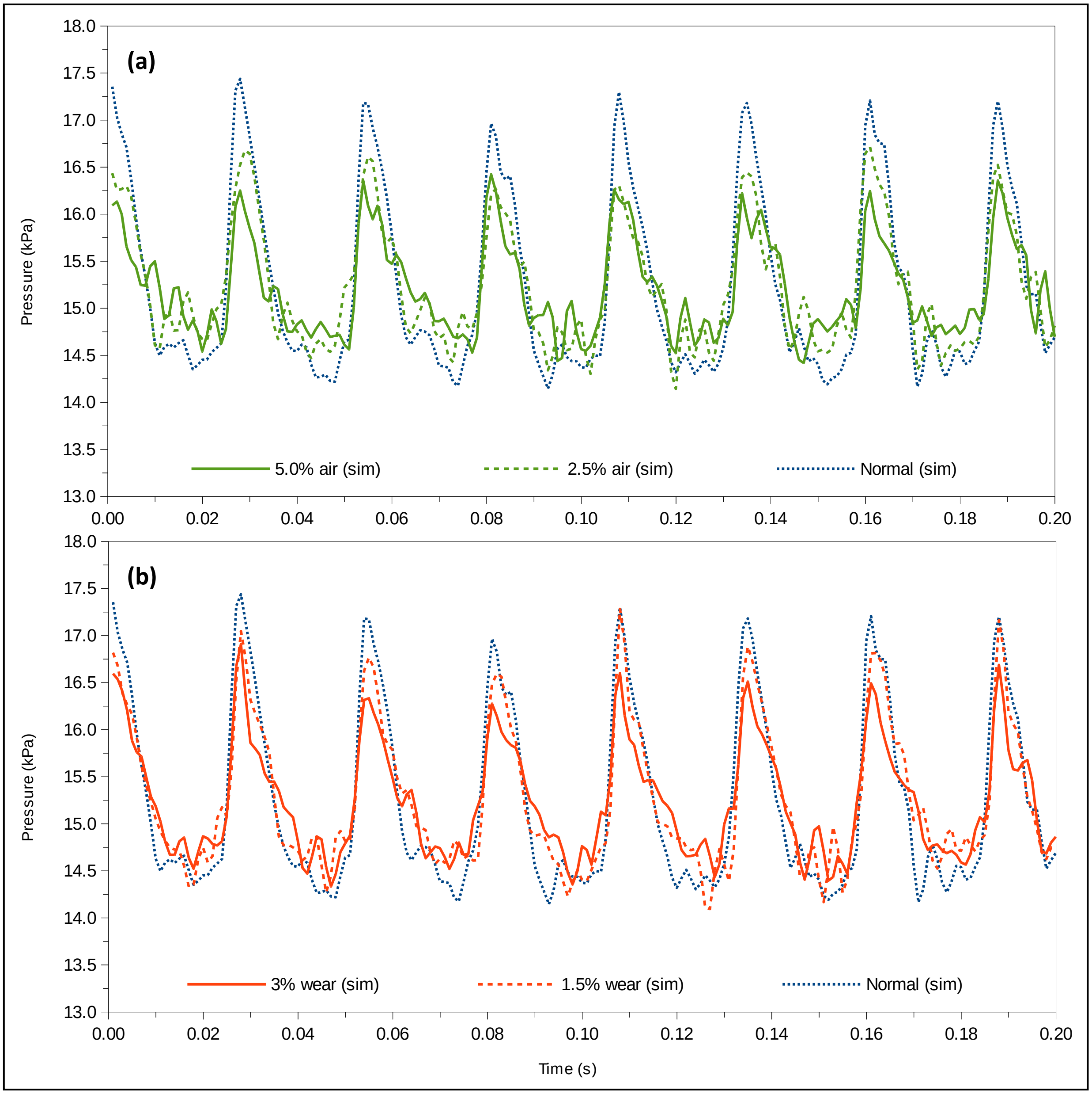


Figure 3.1: a) Simulated pressure signals of a pump at $\boldsymbol{\omega} \leq \mathbf{18.75}\boldsymbol{Hz}$ with increasing gas entrainment; b) Simulated pressure signals of a pump at $\boldsymbol{\omega} \leq \mathbf{18.75}\boldsymbol{Hz}$ with increasing impeller wear.

3.2 Signal Decomposition

To quantify the signal differences resulting from changing severity of gas entrainment or impeller wear, each $\boldsymbol{P}_{\mathbf{out}}$ is decomposed into a set of characteristic statistical values called features. Each feature is determined solely through numerical manipulation of $\boldsymbol{P}_{\text{out}}$. These features form the group of inputs used to classify air entrainment or impeller wear at a given state.

The proposed features for both phenomena are discussed in parallel below, grouped according to the domain in which they are extrapolated from. To permit interpretation and validation, only features with justifiable physical significance[6] are considered in the analysis. Ideally, each feature is expected to exhibit a meaningful change in the signals across the range of measured states. Each description contains the calculation method and the behavior it is intended to quantify. Their influence with regards to classification performance (or lack thereof) is validated in *Chapter 4*.

3.2.1 Time Domain Features

(a) Mean: The arithmetic mean of the discharge pressure signal

$$\overline{p_{\text{out}}} = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{23}$$

is calculated by summing each discrete pressure magnitude and dividing by the total number of data points in the measurement. The mean is necessary for two purposes. First, it serves as a feature in its own right, representing the average pressure being generated at the pump discharge, which is a fundamental parameter of pump operation that should be accounted for. Second, it is employed in calculations of other features that require removal of the steady-state component (i.e. the DC bias) of $\boldsymbol{P}_{\text{out}}$. Each "unbiased" pressure magnitude

$$\hat{x}_n = x_n - \overline{p_{\text{out}}}\,. \tag{24}$$

The unbiased signal is denoted $\widehat{\boldsymbol{P}}_{\text{out}}$. Where it is additionally necessary to scale $\widehat{\boldsymbol{P}}_{\text{out}}$ to fall between -1 and 1, each normalized data point

$$\tilde{x}_n = \frac{\hat{x}_n}{\max(|\widehat{\boldsymbol{P}}_{\text{out}}|)}\,, \tag{25}$$

where $\max(|\widehat{\boldsymbol{P}}_{\text{out}}|)$is the maximum value of $|\hat{x}_n|$ in $\widehat{\boldsymbol{P}}_{\text{out}}$. The normalized pressure signal is denoted $\widetilde{\boldsymbol{P}}_{\text{out}}$.

[6] It is noted that using features with plausible physical meaning is not strictly necessary for this method. The raw time series data, spectral histogram, or other larger feature sets could also be suitable inputs for machine learning. Here, the decision to investigate a set of targeted statistical measures is intended to a) retain a degree of physical interpretability for validation, b) lessen computational load, and c) advise and promote generalizability for the diagnostic method across a variety centrifugal pumps.

(b) Energy: Signal energy

$$E_s = \sum_{n=1}^{N} |\hat{x}_n|^2 \tag{26}$$

is calculated by summing the squares of the unbiased pressure magnitudes in $\widehat{\boldsymbol{P}}_{\mathrm{out}}$ over the length of the sample. The term "signal energy" is not applied in the physics sense, but by the signal processing convention, where it quantifies the magnitude of the periodic portion of the signal. It is applied to generalize overall amplitude of the signal – "loudness," to use the acoustic analogy – but also correlates to the frequency and waveform shape itself (i.e. a sine wave, square wave, and triangle wave of the equivalent amplitudes and frequencies will have different signal energies). Applying (26) to the pressure signals in the preliminary simulations and normalizing the resulting trends yields the following trends for gas entrainment (Figure 3.2a) and radial impeller wear (Figure 3.2b).

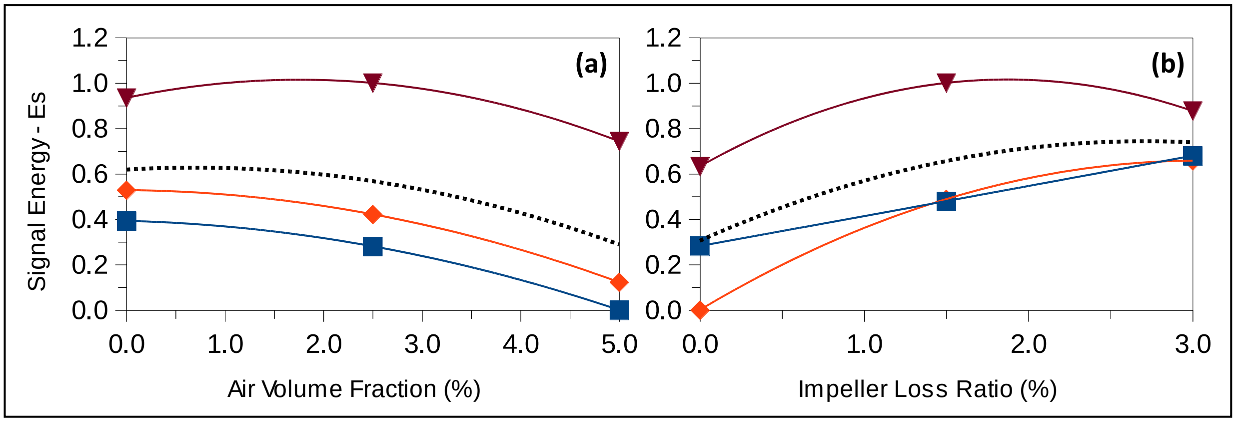


Figure 3.2: $\boldsymbol{E}_s$ as a function of a) $\boldsymbol{\varphi}_{\mathbf{air}}$ and b) $\boldsymbol{\alpha}_{\mathbf{imp}}$. The magnitudes are normalized for comparison. The blue squares, orange diamonds, and maroon inverted triangles correspond to the states with low ($\boldsymbol{\omega} = \mathbf{8.75}\boldsymbol{Hz}$), intermediate ($\boldsymbol{\omega} = \mathbf{13.75}\boldsymbol{Hz}$), and high ($\boldsymbol{\omega} = \mathbf{18.75}\boldsymbol{Hz}$) impeller speeds, respectively. The trend in each speed range is illustrated using a second-order polynomial best-fit line. Additionally, the trend for all the states is indicated by the black dotted line. These trend lines are included to illustrate the cases where the feature has a dual dependence on $\boldsymbol{\omega}$, as well and the target phenomenon. This labeling convention is employed in each of the subsequent feature plots.

The magnitude of E_s tends to decrease slightly as gas entrainment worsens. Here, the influence of impeller speed is also apparent, causing a vertical spread at each condition. This is expected, yet is still undesirable with regards to classification. The correlation is less present at higher impeller speeds, which correspond to the top points (maroon) across Figure 3.2a. In Section 4.1.4, a method is discussed to mitigate the influence from ω. Figure 3.2b suggests a positive correlation between E_s and impeller wear, across impeller speeds.

(c) Variance: The signal variance V_s can be described using central moments, where

$$m_i = \frac{1}{N} \sum_{n=1}^{N} (\hat{x}_n)^i \tag{27}$$

is the i^{th} moment about the signal mean. Variance is defined as the second central moment. Substituting $i = 2$ into (27) yields

$$V_s = m_2 = \frac{1}{N} \sum_{n=1}^{N} (\hat{x}_n)^2. \tag{28}$$

Variance characterizes the overall spread of the data points in the signal about their average value. Applying (28) to the initial simulations yields the trends in Figure 3.3.

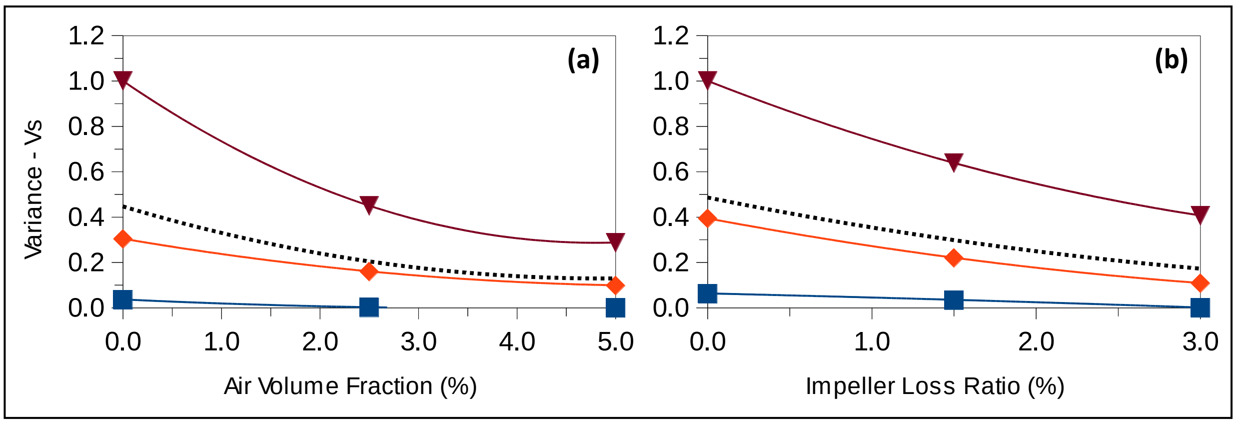


Figure 3.3: $\boldsymbol{V_s}$ as a function of a) $\boldsymbol{\varphi_{\text{air}}}$ and b) $\boldsymbol{\alpha_{\text{imp}}}$. The magnitudes are normalized for comparison.

Both gas entrainment and impeller wear exhibit a decrease in variance as their severity increases. This infers that a reduction in variance may be a useful indication of both conditions. The simulations again demonstrate a secondary correlation between variance of p_{out} and ω, as indicated by the vertical spacing at each respective trend line.

(d) Skewness: The signal skewness G_s is calculated using the Fisher-Pearson coefficient of skewness as the ratio of moments in (29) [69].

$$G_s = \frac{m_3}{{m_2}^{3/2}} \tag{29}$$

Substituting (27) into (29) with the respective values of i and simplifying using (28) yields the calculation for the skewness coefficient.

$$G_\mathrm{s} = \frac{\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^3}{\left[\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^2\right]^{3/2}} = \frac{\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^3}{{V_\mathrm{s}}^{3/2}} \tag{30}$$

Skewness describes the asymmetry of the signal about its mean. Physically, this reflects the relative orientation of the peaks and valleys of the pressure fluctuations during each blade-passing event. A signal with a positive skewness will have the majority of data points fall below the signal average and the reverse for a negative skewness coefficient. Calculating the skewness of the initial simulations generates the trends in Figure 3.4.

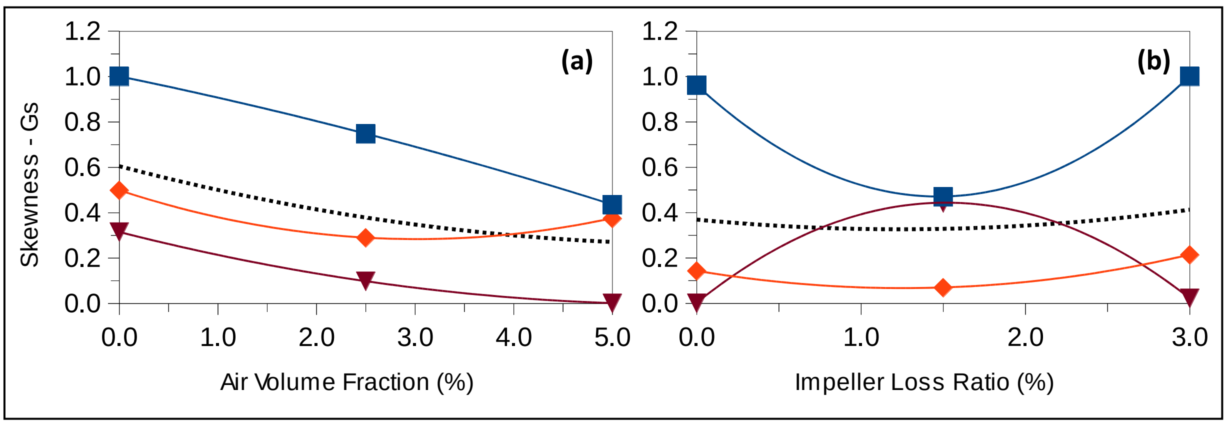


Figure 3.4: $\boldsymbol{G}_\mathbf{s}$ as a function of a) $\boldsymbol{\varphi}_\mathbf{air}$ and b) $\boldsymbol{\alpha}_\mathbf{imp}$. The magnitudes are normalized for comparison.

Figure 3.4a shows a perceptible decrease in $\boldsymbol{G}_\mathbf{s}$, suggesting the sampled pressures tend to increase with respect to their signal mean as φ_air becomes more severe. The same calculation for simulated wear (Figure 3.4b) does not yield an obvious change, signifying that skewness may not be of particular value in characterizing impeller wear.

(e) Kurtosis: Signal kurtosis K_s is calculated using the Fisher[7] convention [69] as the ratio of moments

[7] The alternative kurtosis calculation method is the Pearson convention, which is identical to (45), but with 3 subtracted from the right hand side. This corresponds to removing the kurtosis of a standard normal distribution, leaving the "excess" kurtosis, which

$$K_s = \frac{m_4}{{m_2}^2}\,. \tag{31}$$

Substituting (27) into (31) with the respective values of i and simplifying using (28) yields

$$K_s = \frac{\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^4}{\left[\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^2\right]^2} = \frac{\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_n)^4}{{V_s}^2}\,. \tag{32}$$

The magnitude of K_s describes the tendency of the signal to contain data points far removed from its mean. Physically, it is correlated to the magnitudes and breadths of the pressure peaks and valleys. The kurtosis values for the preliminary simulations are shown in Figure 3.5.

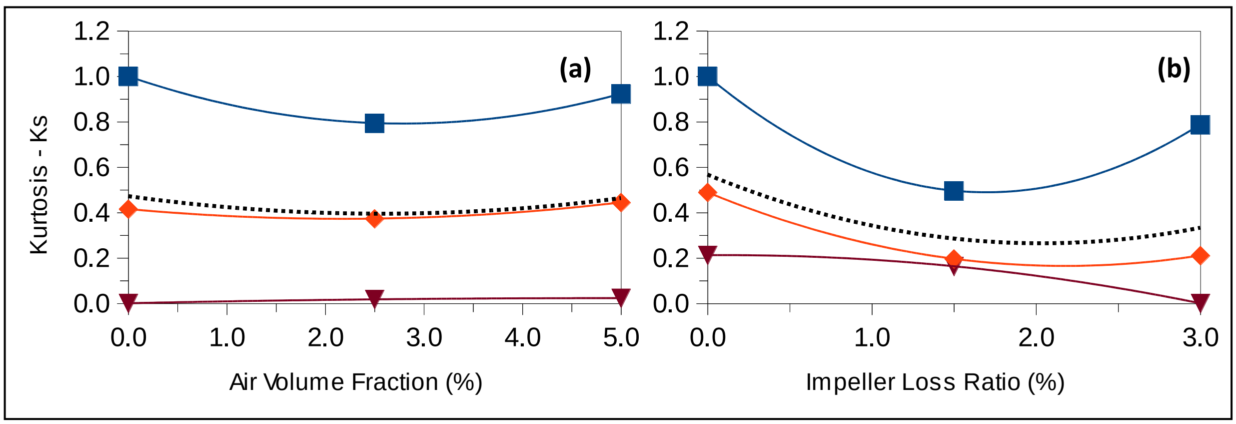


Figure 3.5: $\boldsymbol{K_s}$ as a function of a) $\boldsymbol{\varphi_{\mathrm{air}}}$ and b) $\boldsymbol{\alpha_{\mathrm{imp}}}$. The magnitudes are normalized for comparison.

The preliminary states do not exhibit a strong correlation between kurtosis and the target conditions. However, experimental pressure measurements reported by T. Xie [48] indicate that severe entrainment can introduce broadband fluctuations, which would influence the signal shape dramatically, affecting the kurtosis. It is plausible that the lack of correlation in Figure 3.5 is the byproduct of the trend being concealed by the stronger correlation to impeller speed, but more data is necessary to make that determination. This is evaluated further in *Chapter 4*.

(f) Autocorrelation Energy: Autocorrelation is a method for quantifying the periodicity of a signal. The discrete autocorrelation coefficients R_l of the signal are calculated by correlating $\widetilde{\boldsymbol{P}}_{\mathrm{out}}$ to a duplicate

can be useful for characterizing signals in some cases. However, in this research, each feature is ultimately scaled to fall between 0 and 1, so the inclusion of the -3 term does not provide any benefit.

of itself $\widetilde{\boldsymbol{P}}'_{\text{out}}$, which is identical, but incorporates a time-shift l. The correlation is equal to dot product of the signals at the given time shift, or

$$R_l = \widetilde{\boldsymbol{P}}_{\text{out}} \cdot \widetilde{\boldsymbol{P}}'_{\text{out}} = \sum_{n=1}^{N} \tilde{x}_n \tilde{x}_{n+l} \,. \tag{33}$$

Computing R_l for $0 \leq l \leq N$ at steps equal to t_{s}, assembling the coefficients into an array, and normalizing[8] yields the autocorrelation signal $\boldsymbol{R}_{\text{s}}$.

For an unbiased periodic signal, $\boldsymbol{R}_{\text{s}}$ will oscillate between positive and negative correlation, decaying as the time overlap between the signals shrinks. The addition of noise will hasten the decay. Calculating the energy of $\boldsymbol{R}_{\text{s}}$ using the same method as (26) yields the autocorrelation energy E_{R}. The autocorrelation signals for the preliminary simulations are shown in Figure 3.6.

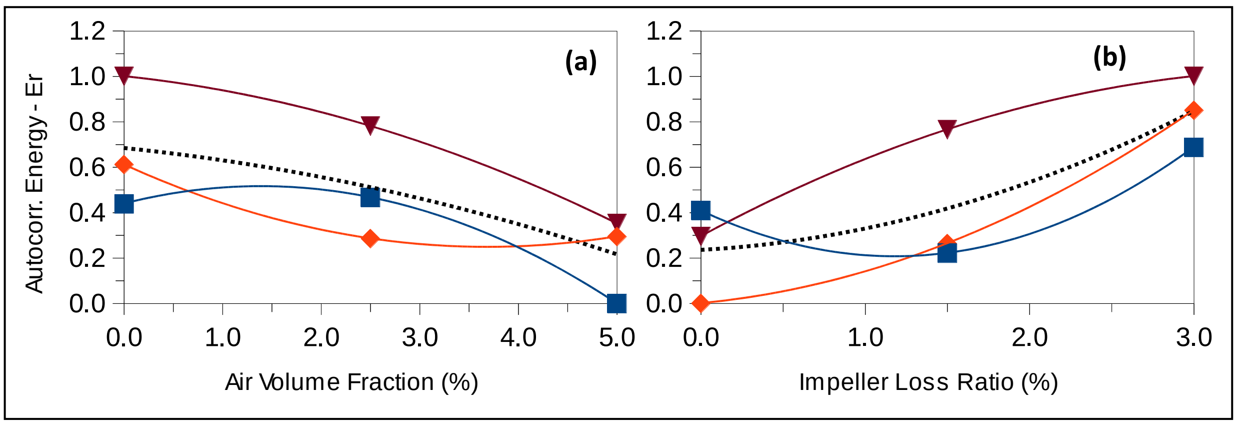


Figure 3.6: $\boldsymbol{E}_{\mathbf{R}}$ as a function of a) $\boldsymbol{\varphi}_{\mathbf{air}}$ and b) $\boldsymbol{\alpha}_{\mathbf{imp}}$. The magnitudes are normalized for comparison.

It is observed in Figure 3.6a that the addition of air to the fluid decreases the autocorrelation, which indicates that the introduction of air causes the periodicity of p_{out} to degrade. This is supported by the experimental measurements in Figure 1.10. Increasing radial impeller wear creates the opposite effect, as shown in Figure 3.6b. The eroding impeller diameter increases autocorrelation energy, suggesting the fluctuations are more periodic. This is notable because it implies that E_{R} may be an effective feature to differentiate between wear and air entrainment.

---

[8] The autocorrelation signal is normalized by dividing each discrete coefficient by the maximum value in the array. This maximum value occurs at the first value, where the time-shift $l = 0$. Here the correlation is defined to be 1 because the signals $\widetilde{\boldsymbol{P}}_{out}$ and $\widetilde{\boldsymbol{P}'}_{out}$ are equivalent.

3.2.2 Frequency Domain Features

Features from the frequency domain are also employed. Each signal is converted to a spectrum in the frequency domain using the Fast Fourier Transform (FFT) algorithm. This spectrum $\boldsymbol{X}$ comprises a list of magnitudes X_m, each at a corresponding discrete frequency f_m with index m. The frequency spectrum has a total of M data points. The peak frequency in the spectrum $f_M = f_{\mathrm{nyq}}$, where $f_{\mathrm{nyq}} = (f_s/2)$ – the Nyquist frequency. However, since the raw signals are filtered using a 400 Hz LPF prior to analysis, frequencies beyond 400 Hz are not considered, even though they fall below f_{nyq}.

Studies by Yuan [42] and Suhane [43] show that the dominant frequency peak in the discharge pressure signal is the BPF. Noting this, the BPF can inversely be determined by extracting the frequency corresponding to the peak magnitude in $\boldsymbol{X}$. This is critically important because it permits use of the BPF in the analysis and classification. Its determination requires no external input beyond the dynamic pressure measurement. Were it otherwise, its incorporation would not be justified in this method. Normalizing the frequencies in $\boldsymbol{X}$ to the BPF allows to content of each spectrum to be compared independent of its individual RF/BPF, simplifying the analysis.

(g) Energy: The total energy of the frequency spectrum E_{FFT} is calculated using a method similar to (26). However, instead of the unbiased data points, the calculation is done using the magnitudes from the frequency spectrum, with each frequency scaled to a multiple of the BPF, with the spectral peak at the BPF normalized to 1.

$$E_{\mathrm{FFT}} = \sum_{m=1}^{M} |X_m|^2 \tag{34}$$

Parseval's theorem states that the energy of a signal is equal to the energy of its frequency spectrum (being a linear transformation). However, by the scaling the frequencies using the BPF of the respective signal, then normalizing against this peak magnitude, the linearity is not upheld. Here, as a result, E_{FFT} characterizes the minor spectral peaks with respect to the main frequency peak, rather than the total energy of the signal itself. It can be influenced by a number of factors, including the noise floor, growth or decay at specific fluctuation frequencies, and changing waveforms (which can add or detract frequency peaks at their associated harmonics).

Applying (34) to the preliminary simulations yields the trends shown in Figure 3.7.

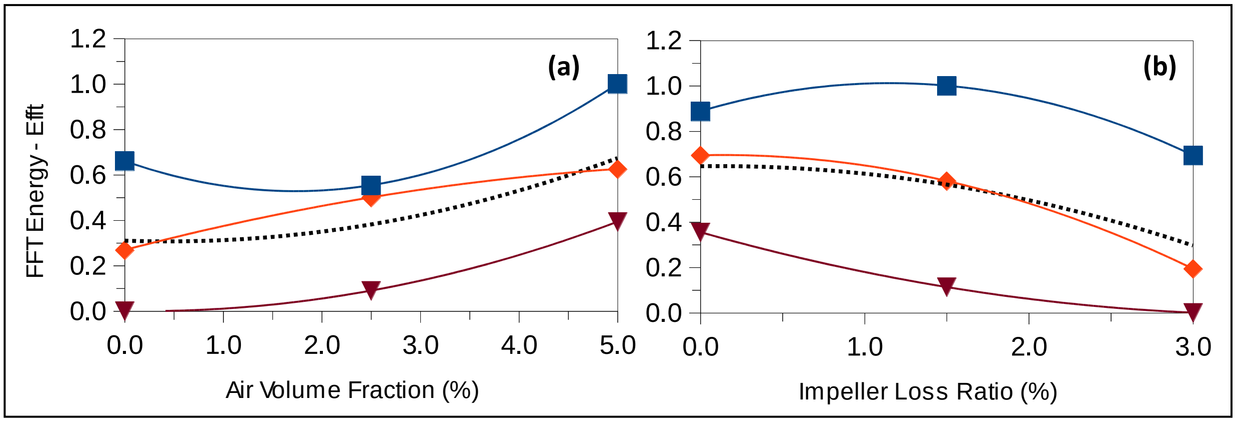


Figure 3.7: $\boldsymbol{E}_{\mathbf{FFT}}$ as a function of a) $\boldsymbol{\varphi}_{\mathbf{air}}$ and b) $\boldsymbol{\alpha}_{\mathbf{imp}}$. The magnitudes are normalized for comparison.

The simulations demonstrate that E_{FFT} has slight positive and negative correlations to φ_{air} and α_{imp}, respectively. However, the vertical spread at each condition, caused by the increasing impeller speed, obscures the trend. A larger pool of measurements is needed to more firmly corroborate the relationship between E_{FFT} and the target conditions.

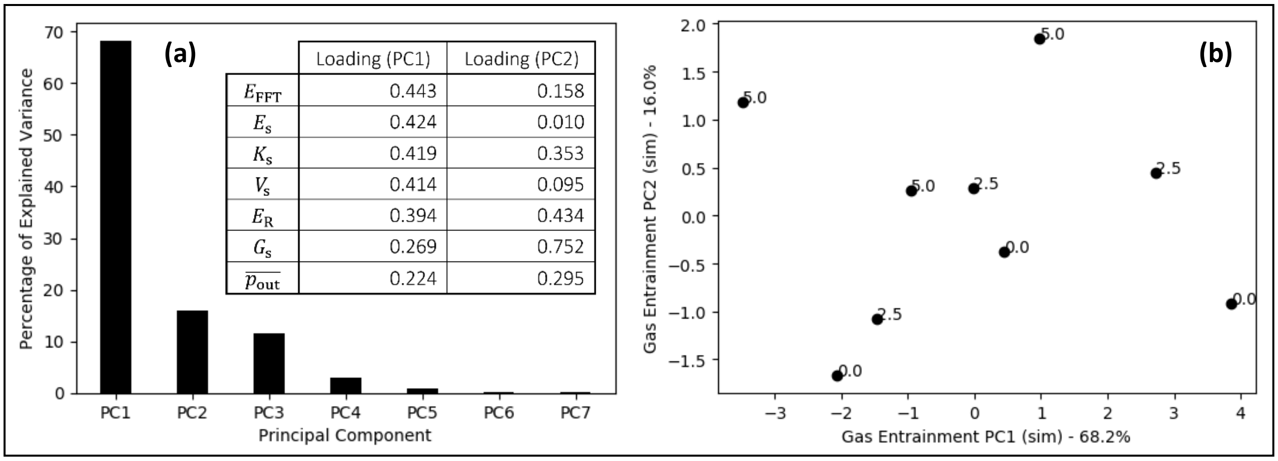
3.3 Feature Space

The seven features from Section 3.2 comprise a feature vector

$$\boldsymbol{F} = [\overline{p_{\mathrm{out}}}, E_{\mathrm{s}}, V_{\mathrm{s}}, G_{\mathrm{s}}, K_{\mathrm{s}}, E_{\mathrm{R}}, E_{\mathrm{FFT}}] \tag{35}$$

for each measured state. The feature vectors are employed as inputs for classification models. Each feature vector is assigned a classification label coinciding with its operating condition – either φ_{air} or α_{imp} – to be referenced as target values during training and testing. The aggregate of the feature vectors across all sampled states forms the feature space $\{\boldsymbol{F}\}$.

Applying principal component analysis (PCA) to $\{\boldsymbol{F}\}$ of the preliminary gas entrainment simulations yields the trends shown in Figure 3.8.



	Loading (PC1)	Loading (PC2)
E_{FFT}	0.443	0.158
E_{s}	0.424	0.010
K_{s}	0.419	0.353
V_{s}	0.414	0.095
E_{R}	0.394	0.434
G_{s}	0.269	0.752
$\overline{p_{\mathrm{out}}}$	0.224	0.295

Figure 3.8: PCA of the preliminary gas entrainment simulations. a) The corresponding scree plot with an inset table of loadings for the first two principal components (PCs). The loadings are the coefficients by which the input features can linearly combined to construct PCs. They can be interpreted as the strength of influence a given feature has on the orientation of a particular PC. The loadings are ordered by their contribution to PC1; b) The distribution of the simulated states across PC1 and PC2. Each state is labeled with the magnitude of $\boldsymbol{\varphi}_{\mathbf{air}}$.

The initial gas entrainment data suggests that there is a principal component (PC) across which 68% of the feature variance manifests (Figure 3.8a, PC1). Plotted across PC1 and PC2, the preliminary states demonstrate a degree of linear separability, with the most severe gas entrainment states trending toward the top left quadrant and the healthy states trending to the opposite.

Applying PCA to $\{\boldsymbol{F}\}$ of the preliminary impeller wear simulations yields the trends shown in Figure 3.9.

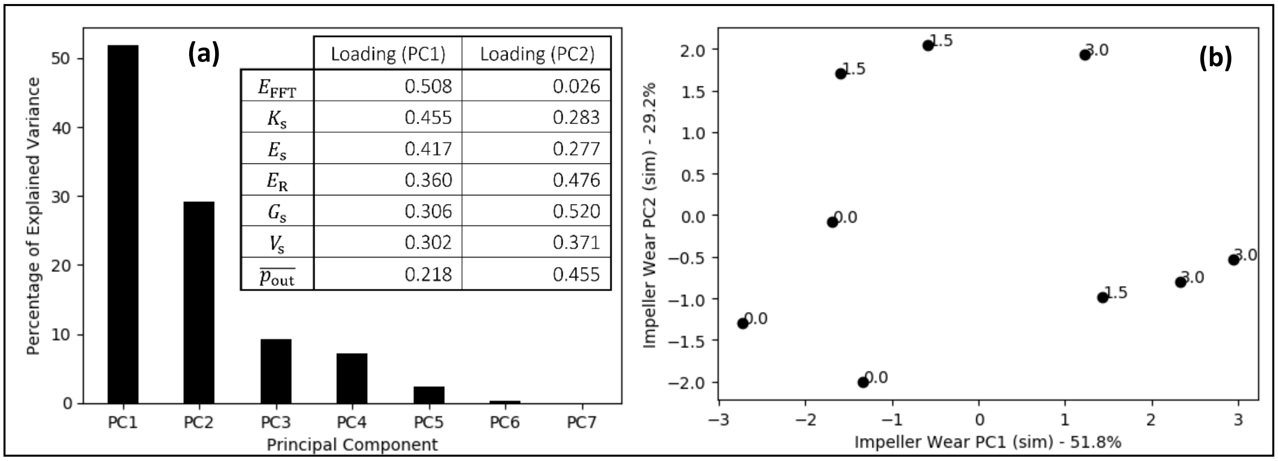


Figure 3.9: PCA of the preliminary impeller wear simulations. a) The corresponding scree plot with an inset table of loadings for PC1 and PC2. The loadings are ranked by their contribution to PC1; b) The distribution of the simulated states across the first two PCs. Each state is labeled with magnitude of $\boldsymbol{\alpha}_{\mathbf{imp}}$.

The initial impeller wear data demonstrates that 52% of the feature variance manifests across PC1 (Figure 3.9a). Correlating PC1 and PC2, Figure 3.9b shows a relatively good linear separability in the target condition, with the most severe wear states trending toward the top right quadrant and the healthy states trending to the opposite.

Figures 3.8 and 3.9 show that the both conditions may be linearly separable using the proposed feature vector. However, this assessment is based on a limited number of states and must be considered as an estimation until it can be validated on a more extensive pool of conditions. Therefore, in selecting prediction models for binary, multi-class, and continuous-value regression, it is assumed that the target conditions will not be perfectly linearly separable for all conditions.

The preliminary features discussed in Section 3.2 reveal characteristic trends across the feature space for both gas entrainment and impeller wear. The correlations between the features and target conditions suggests that classification can be achieved through the use of conventional supervised machine learning methods, in lieu of more complex deep learning methods.

3.4 Classification Models

Three methods are employed to categorize gas entrainment and radial impeller wear; binary (i.e. two-class), multi-class, and continuous-value classification. Each makes a successively more specific prediction of the target condition. The respective learning models employ supervised learning to classify the target phenomena. Each model is developed in Python 3.5. The binary and multi-class neural network algorithms are programmed largely from scratch, whereas the support vector machines (SVM) and random forest regression methods are implemented using the *scikit-learn* library [70].

Binary Classification

The binary classification model predicts the severity of the gas entrainment and impeller wear with respect to acceptability thresholds τ_{air} and τ_{imp}, respectively. Below the threshold, the centrifugal pump is assumed to be operating in a satisfactory condition. The nominal thresholds for gas entrainment and impeller wear are selected to be $\tau_{\mathrm{air}} = 2.0\%$ and $\tau_{\mathrm{imp}} = 1.5\%$, respectively. The binary classification target label "0" is appended to the feature vectors for states that fall below their respective severity threshold. For states that have reached or exceeded their respective threshold, the condition is regarded as degraded and the target label "1" is instead appended.

A multilayer perceptron (MLP) is applied for binary classification. In this application, an MLP with a single hidden layer is used. The structure is shown in Figure 3.10.

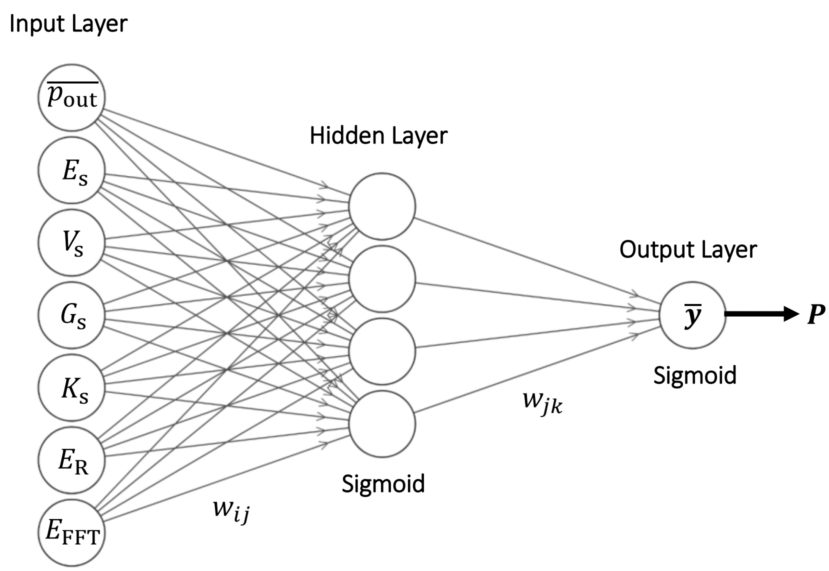


Figure 3.10: Schematic of the MLP architecture used for two-class classification [71].

The model comprises three layers; and input layer, one hidden layer, and the output. The input layer has seven nodes, each corresponding to a feature in $\boldsymbol{F}$. The hidden layer contains four nodes.[9] The output layer is a single node, corresponding to the binary prediction.

In the feedforward stage, the features from a randomly selected state in $\{\boldsymbol{F}\}$ are input. The nodes are connected by associated weights w_{ij} and w_{jk}, with indices i, j, and k corresponding to the input, hidden and output node index, respectively. The weights are assigned random values between -1 and 1, initially. The nodes in the hidden layer compute a weighted average of the incoming values, and output a local prediction using a sigmoid activation function, which tends to 0 for very negative numbers and 1 for very positive numbers. The predictions from the hidden nodes are then weighted again and passed to the output node, which computes a weighted average $\overline{\boldsymbol{y}}$ and generates a final prediction $\boldsymbol{P}$ based on the sigmoid function applied to $\overline{\boldsymbol{y}}$.

The model learns by backpropagating the prediction error and adjusting the network weights. This is done using gradient descent of the mean squared error (MSE)[10] loss function of the prediction, in which the weights are shifted in the direction opposite their respective gradient to find a local minimum that minimizes the MSE (i.e. a least mean squares (LMS) algorithm). The MSE of target prediction Γ_m at index m is calculated as

$$\mathrm{MSE} = \frac{1}{M}\sum_{m=1}^{M}(\Gamma_m - \boldsymbol{P}_m)^2 \, , \tag{36}$$

where M is the total number of output predictions (which is equivalent to the number of nodes in the output layer). When $M = 1$, as in the model in Figure 3.10, (36) simplifies to $MSE = (\Gamma - P)^2$. As the MSE reduces with successive iterations, the values of the weights converge, yielding the configured prediction model. The rate at which the weights are adjusted is controlled by the dimensionless tunable learning rate parameter $\eta_{LR} = 0.02$. The accuracy of the model over training iterations is determined by rounding the

---

9 An appropriate number of nodes in the hidden layer is initially determined through trial and error using an estimated data set, fabricated from the trends of the preliminary simulations. It is later validated as the optimum using experimental data.

10 Ordinarily, MSE would be used as a loss function for regression problems, not classification. However, its application here is warranted because gas entrainment and impeller wear are fundamentally regression-like phenomena, even if the predictions are eventually grouped, using a threshold, into a particular class. C. Beletes *et al.* argue that this is particularly justified in classification problems where some states may be simply be ambiguous based on their features [72].

predictions greater than or equal to 0.5 up to 1 and outputs less than 0.5 to 0, then comparing the rounded prediction to the target classification labels representing the true condition.

For contrast and validation of the MLP's performance, a two-class, soft-margin SVM model is employed in parallel. The applied SVM algorithm uses a polynomial kernel, with a regularization parameter $C_{SVM} = 5$ and kernel coefficient $\gamma_{\mathrm{SVM}} = 5$.

Multi-class Classification

The multi-class classification model categorizes the target phenomenon into a group called a class, which corresponds to a designated severity range. Gas entrainment conditions are divided into $M_{\mathrm{air}} = 6$ classes; $0 \leq \varphi_{\mathrm{air}} < 1\%$, $1\% \leq \varphi_{\mathrm{air}} < 2\%$, etc., up to $\varphi_{\mathrm{air}} > 5\%$. Radial impeller wear is divided into $M_{\mathrm{wear}} = 3$ classes, corresponding to $0 \leq \alpha_{\mathrm{imp}} < 1\%$, $1\% \leq \alpha_{\mathrm{imp}} < 2\%$, and $2\% \leq \alpha_{\mathrm{imp}}$. For each input state, $\boldsymbol{F}$ is appended with a 1xM array of labels, with each value corresponding to a severity class, from least to most severe. This label array is one-hot encoded, with all values equal to zero, except for the entry corresponding to the true severity class, which is 1 (for example: the target label for $\varphi_{\mathrm{air}} = 0.6\%$ would be $[1\ 0\ 0\ 0\ 0\ 0]$ and the label for $\alpha_{\mathrm{imp}} = 2.6\%$ would be $[0\ 0\ 1]$). These labels are used as the target when calculating the prediction accuracy

The MLP shown in Figure 3.10 can be reconfigured to suit multi-class classification. The structure is shown in Figure 3.11.

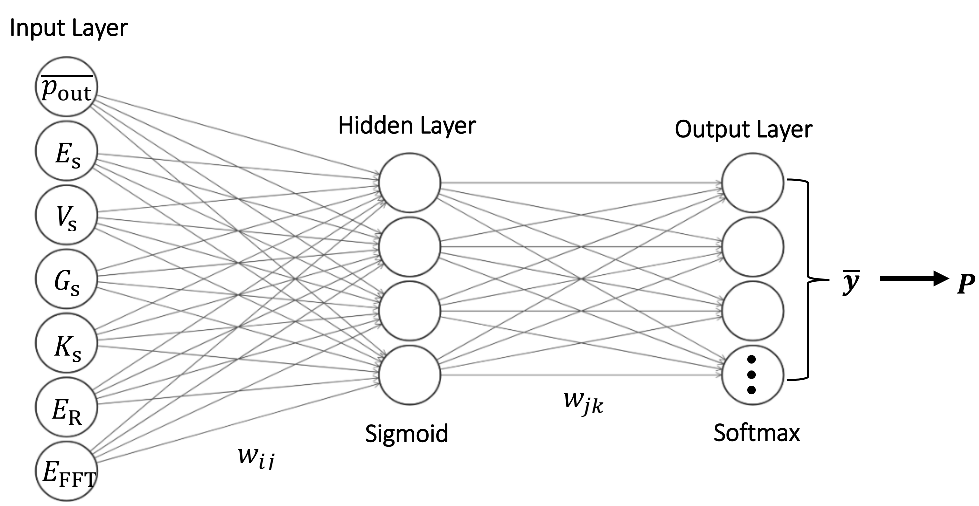


Figure 3.11: Schematic of MLP architecture used for multi-class classification [71]. The gas entrainment model has six output classes. For impeller wear, three classes are used.

The multi-class MLP has M output nodes, one for each severity class. As with the binary classification model, the hidden layer uses the sigmoid activation function. However, here the softmax function

$$\mathrm{softmax}(\bar{\boldsymbol{y}}_k) = \frac{e^{\bar{y}_k}}{\sum_{m=1}^{M} e^{\bar{y}_m}}, \tag{37}$$

is applied to $\bar{\boldsymbol{y}}$ at each output node index k, yielding $\boldsymbol{P}$, a 1xM array of real valued probabilities for each severity class, the sum of which is 1. The element of $\boldsymbol{P}$ with the highest value corresponds to the most likely class. As with the two-class prediction model, gradient descent is used during backpropagation, adjusting the weights so as to minimize the MSE loss function in (36).

As with the binary case, the multi-class MLP is contrasted to a multi-class SVM in parallel. A polynomial kernel is employed again, with $C_{\mathrm{SVM}} = 5$. The kernel coefficient γ_{SVM} is increased to 20 to provide better compensation for potential non-linearities between classes.

Regression

In the final classification step, a random forest algorithm is employed to generate a real-valued prediction of the severity of the target phenomena. The random forest algorithm creates a series of regression trees containing a randomly sampled ("bootstrapped") subset of the training data set. In doing so, the risk of overfitting is curtailed. In this model, the random forest is populated with 10^3 decision trees. The bootstrapped data sets are sampled from the entire feature space, rather than a subset.

4. RESULTS AND DISCUSSION

In this chapter, we present the results and analysis from the gas entrainment and radial impeller wear studies. The entrainment study employs experimental measurements, whereas the wear study uses simulated conditions. In each, we discuss the characteristics of the dynamic pressure measurements, evaluate the relevance of the statistical features put forth in Section 3.2, and demonstrate classification of the target conditions. An analysis of measurement and classification errors follows, after which we propose a method for refining the feature sets to improve classification accuracy.

4.1 Gas Entrainment Study

Dynamic pressure measurements of the discharge flow are collected at 310 states of varying air entrainment, ranging from normal operation to $\varphi_{\mathrm{air}} = 6.0\%$. The samples are divided equally into two sets of 155 states, one for training the learning algorithms and another for testing, with a similar variety of severities in each. The values of ω and φ_{air} for each training state are shown in Table 4.1. The states are indexed by row and column for reference.

Table 4.1: Experimental Gas Entrainment States (Training)

		A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
	ω (Hz) →	**8.9**	**9.4**	**9.9**	**10.4**	**10.9**	**11.4**	**11.9**	**12.4**	**12.9**	**13.4**	**13.9**	**14.4**	**14.9**	**15.4**	**15.9**	**16.4**	**16.9**
1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2		1.6	1.6	1.5	1.4	1.3	1.3	1.2	1.2	1.1	1.1	1.0	1.0	1.0	0.9	0.9	0.9	0.9
3		2.5	2.3	2.2	2.1	2.0	1.9	1.8	1.8	1.7	1.6	1.6	1.5	1.5	1.4	1.4	1.3	1.3
4		3.3	3.1	3.0	2.8	2.7	2.6	2.5	2.4	2.3	2.2	2.1	2.0	2.0	1.9	1.8	1.8	1.7
5		4.2	3.9	3.7	3.5	3.4	3.3	3.1	3.0	2.8	2.7	2.6	2.5	2.4	2.4	2.3	2.2	2.1
6	**φ_{air}**	5.3	4.6	4.4	4.2	4.0	3.8	3.6	3.5	3.3	3.2	3.1	2.9	2.8	2.7	2.7	2.6	2.5
7	(%)→			5.3	5.1	4.8	4.6	4.3	4.2	4.0	3.9	3.7	3.6	3.4	3.3	3.2	3.1	3.0
8							5.2	5.0	4.8	4.6	4.5	4.2	4.1	3.9	3.8	3.7	3.6	3.4
9										5.2	5.0	4.8	4.6	4.4	4.2	4.1	4.0	3.9
10													5.1	4.9	4.8	4.6	4.5	4.3
11													5.7	5.5	5.3	5.2	4.9	4.8
12														6.0	5.9	5.7	5.4	5.2

Each row corresponds to a fixed value of q_{air}. The states without entries are conditions where the void fraction was great enough to cause the pump to stall, preventing a reliable pressure measurement. This occurrence reinforces the initial assumption that entrained gas void fractions exceeding 5% can be considered exceptionally severe. It also parallels experimental outcomes reported by Schäfer [51] and Stan [52].

The pressure waveforms of the ordinary operating states from the training set (Table 4.1, Row 1) are shown in Figure 4.1.

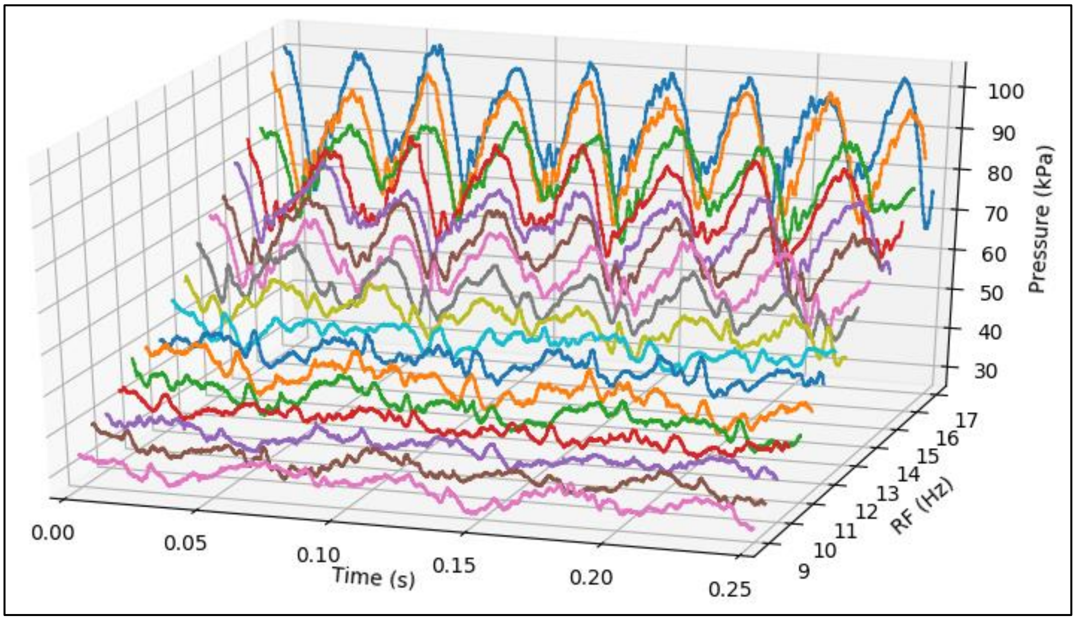


Figure 4.1: Samples of discharge pressure fluctuations across a range of rotating frequencies, during operation without gas entrainment.

As ω increases, so does the discharge pressure. This is the byproduct of the increasing hydraulic resistance caused by faster flow through the system. At the slowest impeller speed, the main oscillating mode of the pressure varies within a band of approximately 7 kPa. At the highest, the fluctuations span 37 kPa. The magnitudes of the higher frequency pressure excursions – those not correlated to blade-passing events – remain consistent across rotating frequencies and are predominantly less than 3 kPa.

The fluctuations in p_{out} also show unexpected alternating depths in the pressure minima following blade-passing events (most perceptible in the highest speed sample in Figure 4.1). This behavior is independent of gas entrainment. It is likely the result of slightly unequal geometry between the two impeller blades. Though it is not within the scope of this work, this phenomenon could potentially be leveraged to detect impeller asymmetry, even if the impeller weight itself is not unbalanced. However, in evaluating the dynamic pressure frequency content, this phenomenon is anticipated to generate a spectral peak at half the BPF (i.e. the RF) that should otherwise not be present.

The inclusion of a substantial void fraction of gas bubbles has a visible influence on the discharge pressure fluctuations. The waveforms of p_{out} at the highest achievable φ_{air} for each rotating frequency (bottom box of each column in Table 4.1) are shown in Figure 4.2.

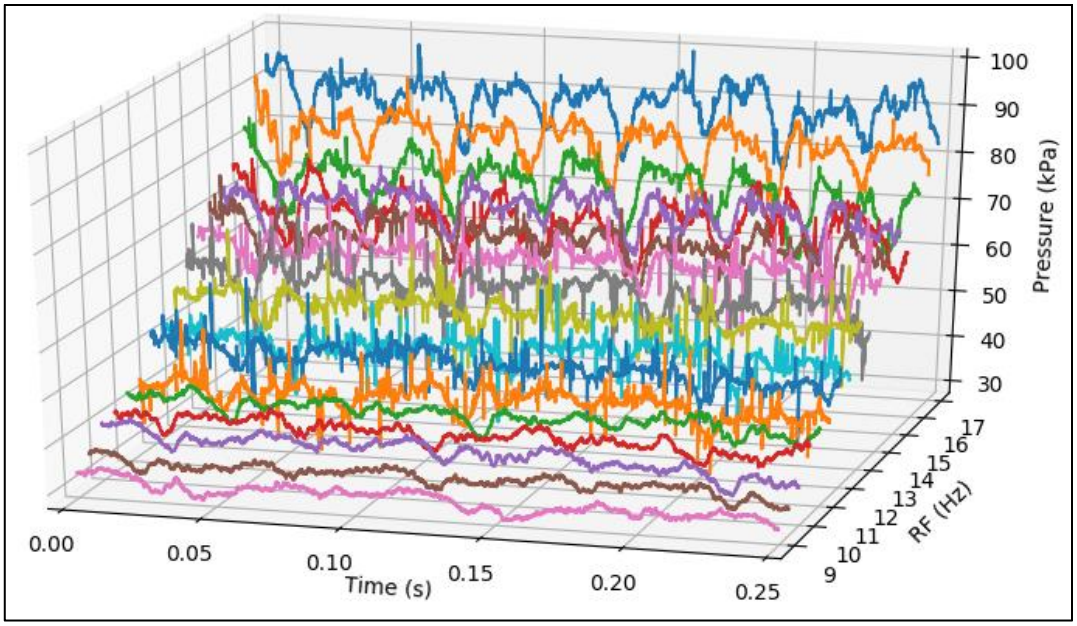


Figure 4.2: Samples of discharge pressure fluctuations across a range of rotating frequencies, each with an air void fraction exceeding 4.5%.

Across frequencies, the main fluctuating mode is not well defined. In all cases, the pressure peaks corresponding to blade passing events are diminished, causing the overall span of the oscillations to decrease. At the lowest speed, the pressure fluctuates within a 5 kPa band, which is approximately 30% narrower than the same margin during normal operating condition. At the highest rotating frequency, the decrease is more substantial. The fluctuation band narrows to 20 kPa, which is a 45% decrease from the normal condition. This reinforces the correlation between variance and the severity of gas entrainment inferred from the initial simulations. The pressure measurements also exhibit an increase in higher frequency fluctuations, which was also demonstrated in the initial simulations. However, there are additional chaotic spikes and drops, particularly in the intermediate rotating frequencies, which were not present in the preliminary simulated measurements. This is plausibly an effect of air bubbles momentarily adhering to the transducer diaphragm then separating or collapsing, similar to behavior reported by

Nakatani [73]. However, validating this hypothesis exceeds the scope of this research. A detail comparison of two discharge pressure signals with varying gas entrainment is shown in Figure 4.3 (states 1M and 11M, from Table 4.1).

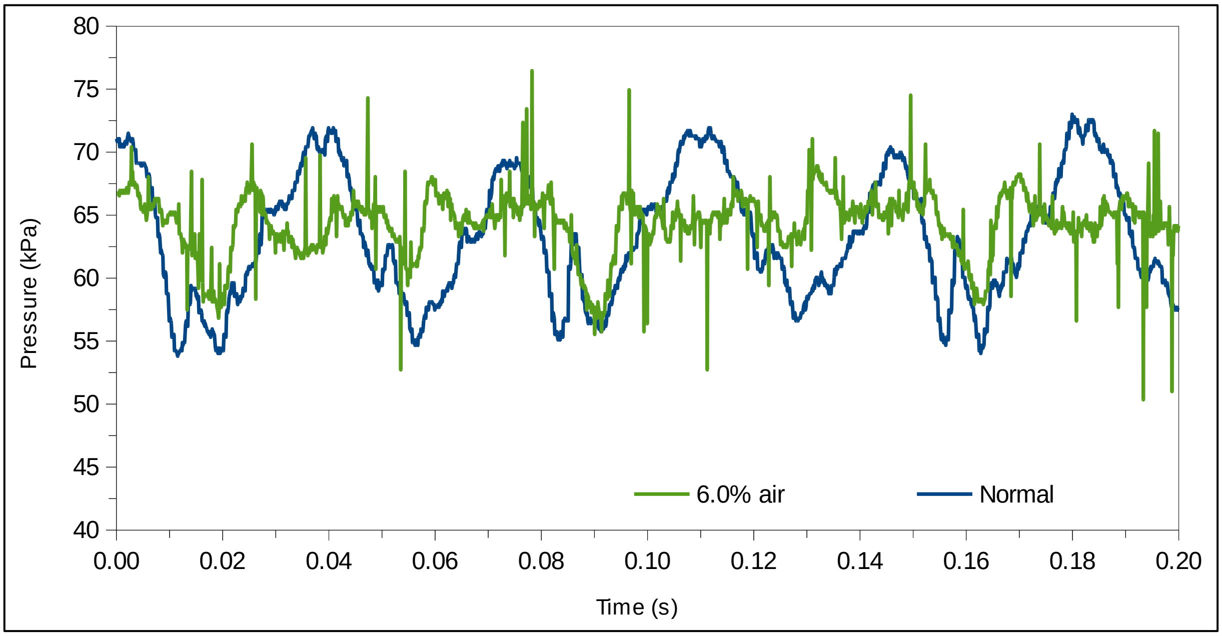


Figure 4.3: The fluctuations of $\boldsymbol{p}_{\text{out}}$ at $\boldsymbol{\omega} = \mathbf{14.9}\boldsymbol{Hz}$ for normal pump operation (blue) and an air entrainment condition with $\boldsymbol{\varphi}_{\text{air}} = \mathbf{6.0}\%$ (green).

At $\varphi_{\text{air}} = 6.0\%$, the main fluctuation mode is substantially degraded. The variance between the local pressure peaks and minima has reduced, and disordered pressure spikes are frequent.

4.1.1 Feature Correlation

Using the methods described in Section 3.2, the measurements of p_{out} are decomposed into statistical features for each state in Table 4.1. The normalized values of each feature are plotted in Figures 4.4a to 4.4g. Each plot also contains four second-order polynomial trend lines corresponding to the features associated with i) the lowest impeller speed, $\omega = 8.9Hz$ (blue line), ii) the median speed, $\omega = 12.9Hz$ (orange line), iii) the highest speed, $\omega = 16.9Hz$ (maroon line), and iv) all the data points (black dashed line). These are included to better illustrate the cases where the correlation between the feature and gas entrainment severity has an additional dependence on ω.

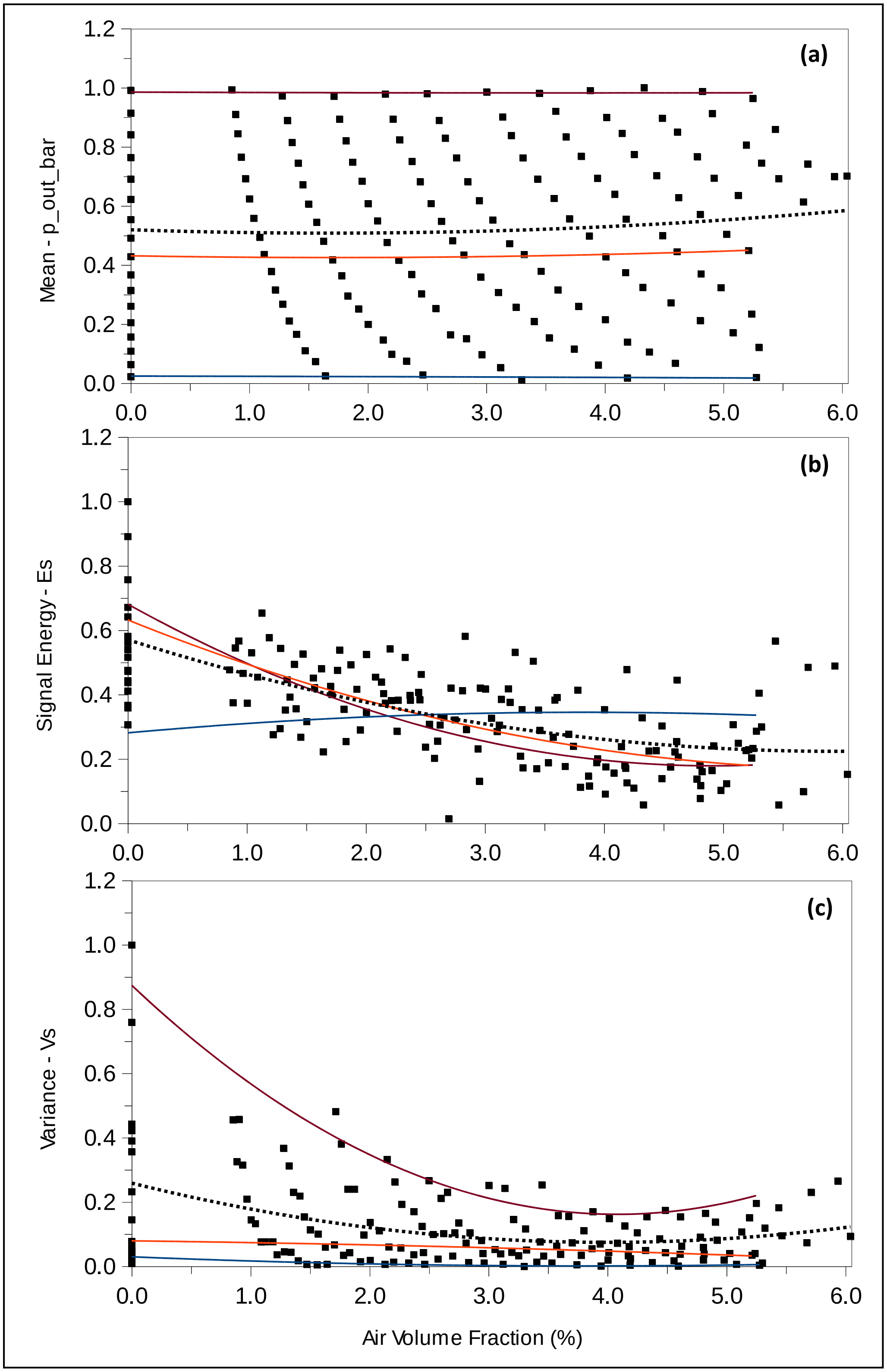


Figure 4.4a-c: The normalized values of a) $\overline{p_{\mathbf{out}}}$, b) $\boldsymbol{E}_{\mathbf{s}}$, and c) $\boldsymbol{V}_{\mathbf{s}}$ as a function of $\boldsymbol{\varphi}_{\mathbf{air}}$.

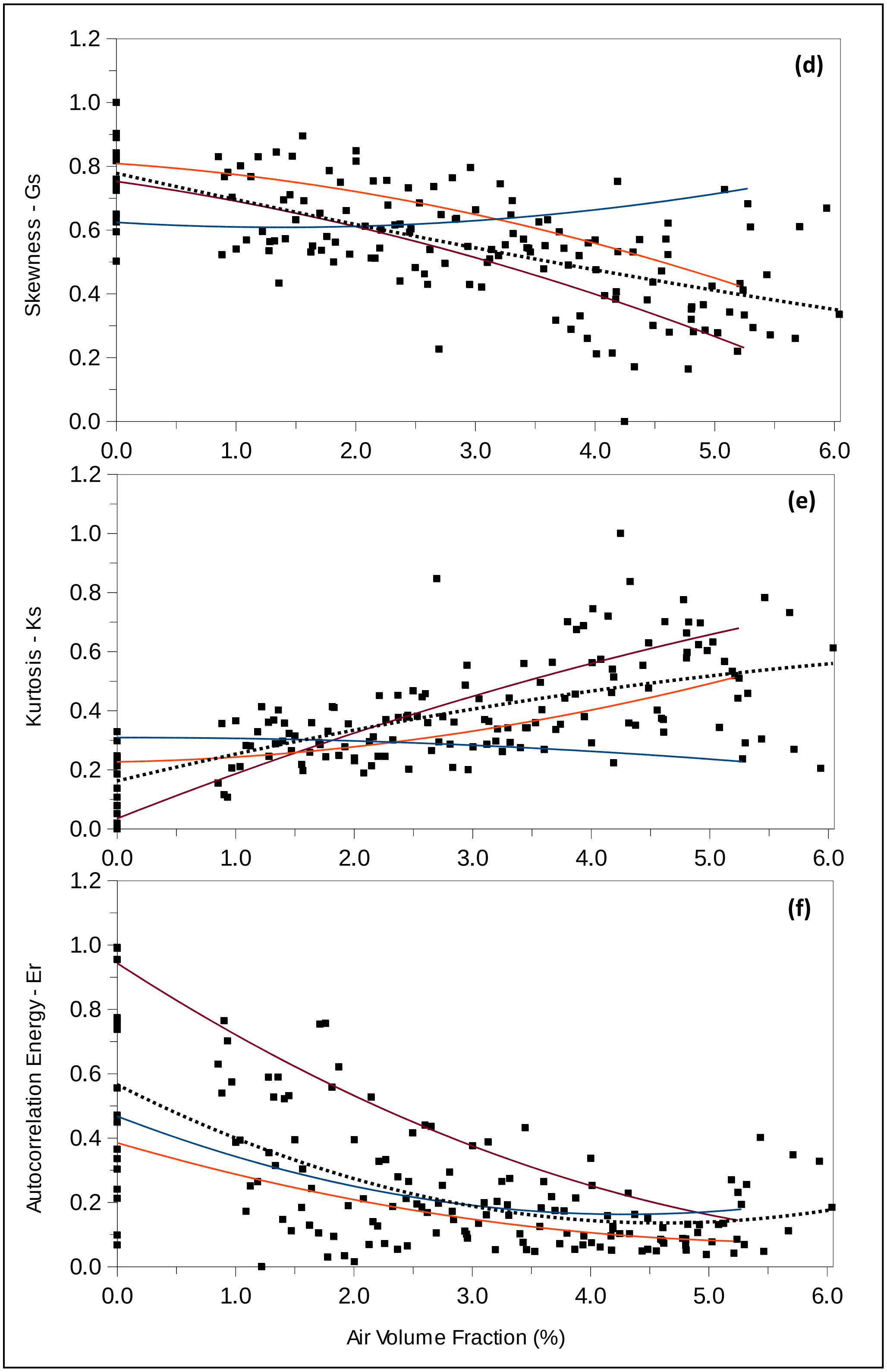


Figure 4.4d-f: The normalized values of d) $\boldsymbol{G_s}$, e) $\boldsymbol{K_s}$, and f) $\boldsymbol{E_R}$ as a function of $\boldsymbol{\varphi_{\mathrm{air}}}$.

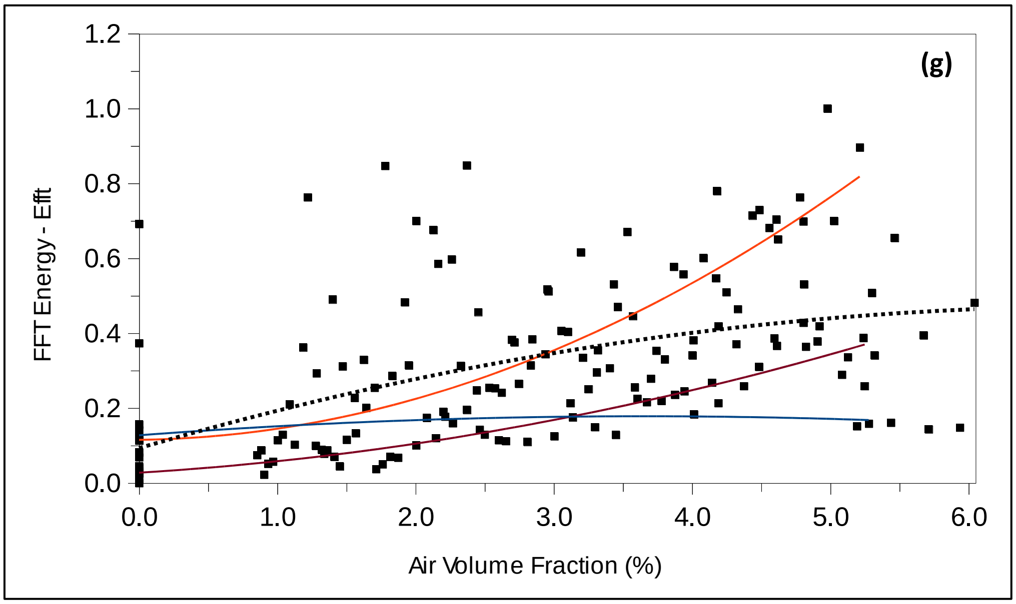


Figure 4.4g: The normalized values of E_{FFT} as a function of φ_{air}.

Figure 4.4a demonstrates that average pressure $\overline{p_{\mathrm{out}}}$ has a poor correlation to gas entrainment for all of the experimental states. Neither the high, median, nor low speed measurements indicate a meaningful positive or negative association with increasing void fraction. Instead, the vertical spacing of the respective trend lines shows that the average discharge pressure is highly correlated to the impeller speed. This suggests that $\overline{p_{\mathrm{out}}}$ may instead function to orient input features with respect to their rotating frequency during classification.

From 4.4a, it can also be inferred that the slopes of the trend lines indicate a correlation to φ_{air} (or, for the case of $\overline{p_{\mathrm{out}}}$, an absence of correlation) and their respective deviations from each other (whether vertical spread, dissimilar slopes, etc.) suggest a dependence on ω.

The samples of the signal energy E_s in Figure 4.4b are an example of mutual dependence on both φ_{air} and ω. At the lowest impeller speed, the E_s does not exhibit any correlation to φ_{air}. However, in both the median and high-speed states, as well as the overall average, E_s decreases. This agrees with the simulated states shown in Figure 3.2a. Physically, this suggests that the average amplitudes of the pressure fluctuations diminish as gas entrainment becomes more severe, but only if the impeller is rotating fast enough to exhibit the effect. If the impeller is moving slowly, the correlation weakens.

Figure 4.4c shows that variance V_s has a negative correlation to φ_{air} at high impeller speeds. This closely parallels the simulated results in Figure 3.3a and agrees with the observations of Figure 4.2. The correlation at intermediate and slow impeller speeds is unclear, due to the dominating influence of the high speed V_{s} values. In Section 4.2.6, a method is discussed for normalizing out the influence of ω, clarifying the relationship.

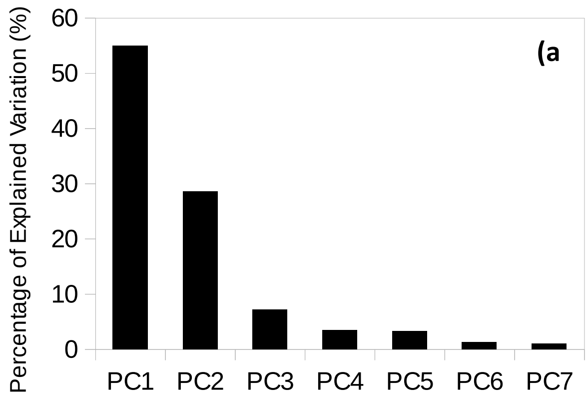
In Figure 4.4d, a negative correlation between skewness G_{s} and φ_{air} is observed at intermediate and high speeds. This matches the trends of the initial simulations in Figure 3.4a. A reduction in G_{s} indicates that the values of p_{out} have decreased with respect to the signal mean. Physically, this means a) the pressure peaks associated with a blade passing event have been diminished, or b) the proceeding pressure drops have been accentuated. The waveforms in Figures 4.1 and 4.2 corroborate the former. The pressure peaks at each blade-passing event are truncated when air is introduced, skewing the measurements downward without significantly affecting the mean. The correlation is less perceptible at the slowest impeller speed. This again suggests that the value of G_{s} as a feature for classification diminishes as the impeller slows.

The relationship between Kurtosis K_{s} and φ_{air} (Figure 4.4e) is comparable to that of G_{s}, though with an opposite sign. Intermediate and high speeds exhibit a positive correlation, signifying that the individual values of p_{out} tend to become focused about the mean as gas entrainment worsens. However, the association is again weakened as the impeller slows.

Figure 4.4f shows that autocorrelation energy E_{R} has a strong negative correlation across impeller speeds, but particularly at the highest speed. This matches the preliminary simulated results from Figure 3.6a and substantiates the observation that increasing φ_{air} contributes disorder to the pressure fluctuations, reducing the periodicity of the measurement. The vertical spacing of the trend lines suggests an additional dependence on ω.

The energy of the FFT spectrum E_{FFT} increases with φ_{air} (Figure 4.4g), owing to the increase in higher frequency pressure fluctuations. This agrees with the simulated trends in Figure 3.7a. The association lessens with reduced impeller speed.

PCA is applied to evaluate the correlations between these seven input features. The scree plot is shown in Figure 4.5a, with the loadings for PC1 and PC2 adjacent. The first two PCs are plotted with respect to each other in Figure 4.5b.



	Loading (PC1)	Loading (PC2)
E_{R}	0.450	0.246
K_{s}	0.423	0.342
V_{s}	0.422	0.232
E_{FFT}	0.398	0.113
G_{s}	0.375	0.400
E_{s}	0.354	0.448
$\overline{p_{\mathrm{out}}}$	0.124	0.629

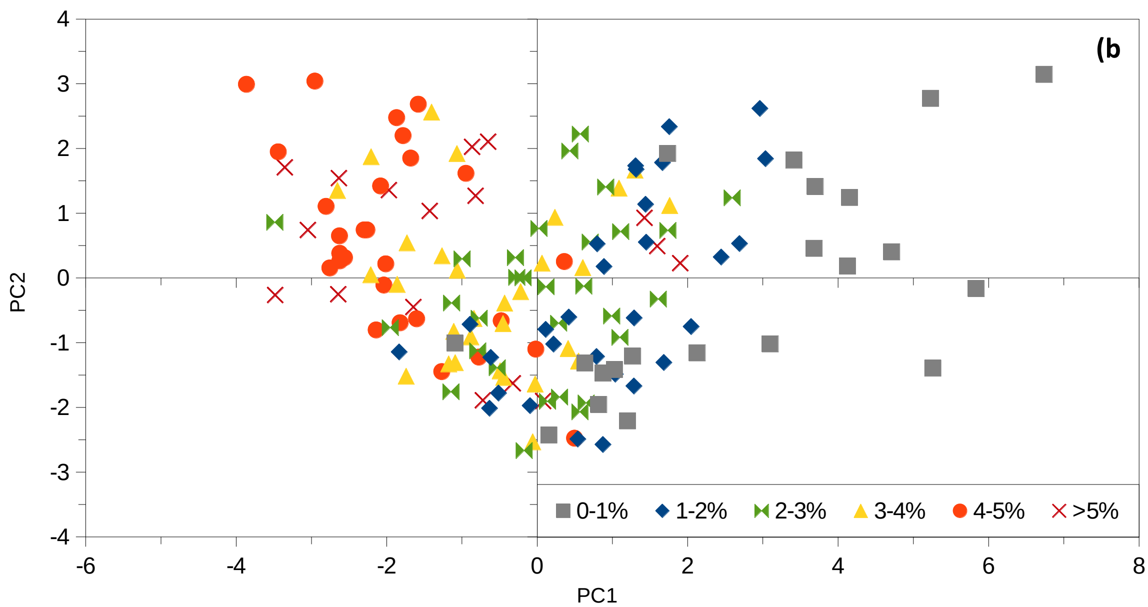


Figure 4.5: a) A scree plot of the PCs of gas entrainment and the loading for each variable with respect to PC1 and PC2, listed in order of their contribution to PC1; b) The distribution of the experimental states across the first two PCs.

In Figure 4.5a, it can be observed that PC1 and PC2 account for more than 80% of the variation across the features. Plotting the experimental states over these two PCs (Figure 4.5b) reveals a degree of grouping, with the states with lower φ_{air} tending to fall to the right, but there are significant regions of overlap. This reinforces the initial assessment that the gas entrainment states should not be treated as linearly separable and justifies the application of non-linear machine learning algorithms for gas entrainment classification.

4.1.2 Binary Classification

To evaluate the performance of the binary classification model, the MLP is trained and tested, independently, 20 times. Training performance is determined by averaging the errors over the final $5 \cdot 10^3$

iterations in each training cycle. With the acceptability threshold for gas entrainment set at $\tau_{\text{air}} = 2.0\%$ the binary classifier correctly predicts the gas entrainment severity category in 90% of the training states.[11] The testing accuracy is 84%. The difference of 6% between training and testing suggests that slight overfitting may be occurring, but the impact is minimal. The binary SVM classifier achieves 88/83% training/testing accuracy using the same states.

The truth table for the binary MLP is shown in Table 4.2.

Table 4.2: Truth Table for Binary MLP Classification of Gas Entrainment (Testing)

True Healthy: **31**	False Healthy: **5**
False Degraded: **20**	True Degraded: **99**

Of the 25 misclassified states in Table 4.2, 80% overestimate the entrainment severity, suggesting the condition had eclipsed the threshold when it has not. Table 4.3 shows the specific conditions misclassified in the testing data set (underlined) by the binary MLP. The true healthy states are shaded gray and the degraded red.

Table 4.3: Gas Entrainment States Misclassified by Binary MLP Model (Testing)

		A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
	ω (Hz) →	**8.9**	**9.4**	**9.9**	**10.4**	**10.9**	**11.4**	**11.9**	**12.4**	**12.9**	**13.4**	**13.9**	**14.4**	**14.9**	**15.4**	**15.9**	**16.4**	**16.9**
1	φ_{air} (%)→	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2		1.4	1.3	1.2	1.2	1.1	1.1	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.7	0.7
3		2.2	2.1	2.0	1.9	1.8	1.7	1.6	1.6	1.5	1.4	1.4	1.3	1.3	1.3	1.2	1.2	1.1
4		3.0	2.9	2.7	2.6	2.5	2.4	2.3	2.2	2.1	2.0	1.9	1.8	1.8	1.7	1.7	1.6	1.6
5		3.9	3.7	3.5	3.3	3.2	3.0	2.9	2.8	2.6	2.5	2.5	2.4	2.3	2.2	2.1	2.1	2.0
6		5.1	4.5	4.3	4.1	3.9	3.7	3.5	3.4	3.2	3.1	3.0	2.9	2.8	2.7	2.6	2.5	2.4
7				5.1	4.9	4.6	4.4	4.1	4.0	3.8	3.7	3.5	3.4	3.3	3.2	3.1	3.0	2.9
8							5.0	4.8	4.6	4.4	4.3	4.0	3.9	3.8	3.7	3.5	3.4	3.3
9										5.0	4.9	4.6	4.5	4.3	4.1	4.0	3.9	3.7
10													5.0	4.8	4.6	4.5	4.3	4.2
11													5.5	5.3	5.2	5.0	4.8	4.7
12														5.9	5.8	5.6	5.3	5.1

[11] The configured MLPs are not perfectly identical each time due to the random starting weights. As such, one or two states may be correctly classified in one configuration and not in another, leading to (negligible) variability in the classification performance.

It is observed that the erroneously classified states are predominantly near the acceptability threshold (2.0%, the gray/red interface), which is to be expected. Only one very severe entrainment condition is misclassified (10Q). The results also suggest that the classification performance degrades at the lowest impeller speeds. This aligns with the observations of the feature trends in Figure 4.4, many of which show weak correlations at the slowest impeller speed. It is proposed that this can be mitigated by normalizing the influence of impeller speed. Though the binary classification model performs reasonably well using the raw feature set, it does not meet the 90% accuracy threshold stated in the research objectives, and therefore must be improved.

4.1.3 Multi-class Classification

The multi-class classification model is evaluated on the same training and testing sets as the binary model. Its performance is likewise evaluated using a similar process. The network is trained and tested 20 separate times, starting from random weights, to determine its accuracy and repeatability. Training accuracy is determined by averaging the prediction errors over the final $5 \cdot 10^3$ iterations in each training cycle. Using classes spanning 1%, the multi-class model places each training state in the correct group with 82% accuracy. The classification accuracy on the testing set is 59%. Of the misclassified testing states, approximately 70% are within one class of the correct prediction. Those that are not are underlined in Table 4.4, with the respective classes color coded.

Table 4.4: Gas Entrainment States Severely Misclassified by Multi-class MLP Model (Testing)

		A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
	ω (Hz) →	**8.9**	**9.4**	**9.9**	**10.4**	**10.9**	**11.4**	**11.9**	**12.4**	**12.9**	**13.4**	**13.9**	**14.4**	**14.9**	**15.4**	**15.9**	**16.4**	**16.9**
1	φ_{air} (%)→	<u>0.0</u>	0.0	<u>0.0</u>	<u>0.0</u>	0.0	0.0	0.0	<u>0.0</u>	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2		1.4	1.3	<u>1.2</u>	1.2	1.1	1.1	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.7	0.7
3		2.2	2.1	2.0	1.9	1.8	1.7	1.6	1.6	1.5	1.4	1.4	1.3	1.3	1.3	1.2	1.2	1.1
4		3.0	2.9	2.7	2.6	2.5	2.4	<u>2.3</u>	2.2	2.1	2.0	1.9	1.8	1.8	1.7	1.7	1.6	1.6
5		<u>3.9</u>	<u>3.7</u>	<u>3.5</u>	<u>3.3</u>	3.2	3.0	2.9	2.8	2.6	2.5	2.5	2.4	2.3	2.2	2.1	2.1	2.0
6		5.1	<u>4.5</u>	4.3	4.1	<u>3.9</u>	3.7	3.5	3.4	3.2	3.1	3.0	2.9	2.8	2.7	2.6	2.5	2.4
7				5.1	4.9	4.6	4.4	4.1	4.0	3.8	3.7	3.5	3.4	<u>3.3</u>	3.2	3.1	3.0	2.9
8							5.0	4.8	<u>4.6</u>	4.4	4.3	4.0	3.9	3.8	3.7	3.5	3.4	3.3
9										5.0	4.9	<u>4.6</u>	4.5	4.3	4.1	4.0	3.9	3.7
10													5.0	4.8	4.6	4.5	<u>4.3</u>	<u>4.2</u>
11													5.5	5.3	5.2	5.0	4.8	4.7
12														5.9	<u>5.8</u>	5.6	5.3	5.1

As with the binary MLP prediction model, it is observed that the majority of the misclassifications occur at the slower impeller speeds, where the correlations between the input features and target condition are weakest. The multi-class SVM model yields training and testing prediction accuracies of 80% and 60%, respectively. These prediction results reinforce the need to improve the relevance of the features for the low speed states.

4.1.4 Regression

The random forest regression model is implemented on the gas entrainment training and testing sets. The median prediction error on the training set is a void fraction within 0.20 of the true φ_{air} percentage value (standard deviation of 0.26). The median prediction error on the testing set is within 0.47 of the true severity (standard deviation of 0.73). Considering the testing predictions miss the true void fraction by more than 1.25, it can be observed that they occur with much greater frequency below $\omega = 12.9Hz$. These states are underlined in Table 4.5.

Table 4.5: Gas Entrainment States Severely Misclassified by Regression Model (Testing)

		A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
	ω (Hz) →	**8.9**	**9.4**	**9.9**	**10.4**	**10.9**	**11.4**	**11.9**	**12.4**	**12.9**	**13.4**	**13.9**	**14.4**	**14.9**	**15.4**	**15.9**	**16.4**	**16.9**
1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2		1.4	1.3	1.2	1.2	1.1	1.1	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.8	0.8	0.7	0.7
3		2.2	2.1	2.0	1.9	1.8	1.7	1.6	1.6	1.5	1.4	1.4	1.3	1.3	1.3	1.2	1.2	1.1
4		3.0	2.9	2.7	2.6	2.5	2.4	2.3	2.2	2.1	2.0	1.9	1.8	1.8	1.7	1.7	1.6	1.6
5		3.9	3.7	3.5	3.3	3.2	3.0	2.9	2.8	2.6	2.5	2.5	2.4	2.3	2.2	2.1	2.1	2.0
6	**φ_{air}**	5.1	4.5	4.3	4.1	3.9	3.7	3.5	3.4	3.2	3.1	3.0	2.9	2.8	2.7	2.6	2.5	2.4
7	**(%)→**			5.1	4.9	4.6	4.4	4.1	4.0	3.8	3.7	3.5	3.4	3.3	3.2	3.1	3.0	2.9
8							5.0	4.8	4.6	4.4	4.3	4.0	3.9	3.8	3.7	3.5	3.4	3.3
9										5.0	4.9	4.6	4.5	4.3	4.1	4.0	3.9	3.7
10													5.0	4.8	4.6	4.5	4.3	4.2
11													5.5	5.3	5.2	5.0	4.8	4.7
12														5.9	5.8	5.6	5.3	5.1

Though the prediction accuracy of the regression model is adequate for an estimate of gas entrainment, this result reiterates that the input features cannot precisely quantify the target condition in low speed states and must be improved.

4.1.5 Feature Optimization

The correlations between the input features and gas entrainment severity are improved by minimizing their respective feature variation caused by impeller speed. The objective is to improve the coefficient of determination R^2 of the input features' trends with respect to gas entrainment. To do so, ω is first extracted from the FFT spectrum at each state by isolating the frequency corresponding to the highest peak (the BPF) and dividing that value by $N_{\mathbf{imp}}$.

Plotting each input feature against its ω values reveals its association to impeller speed. Figure 4.6 demonstrates this correlation for V_s.

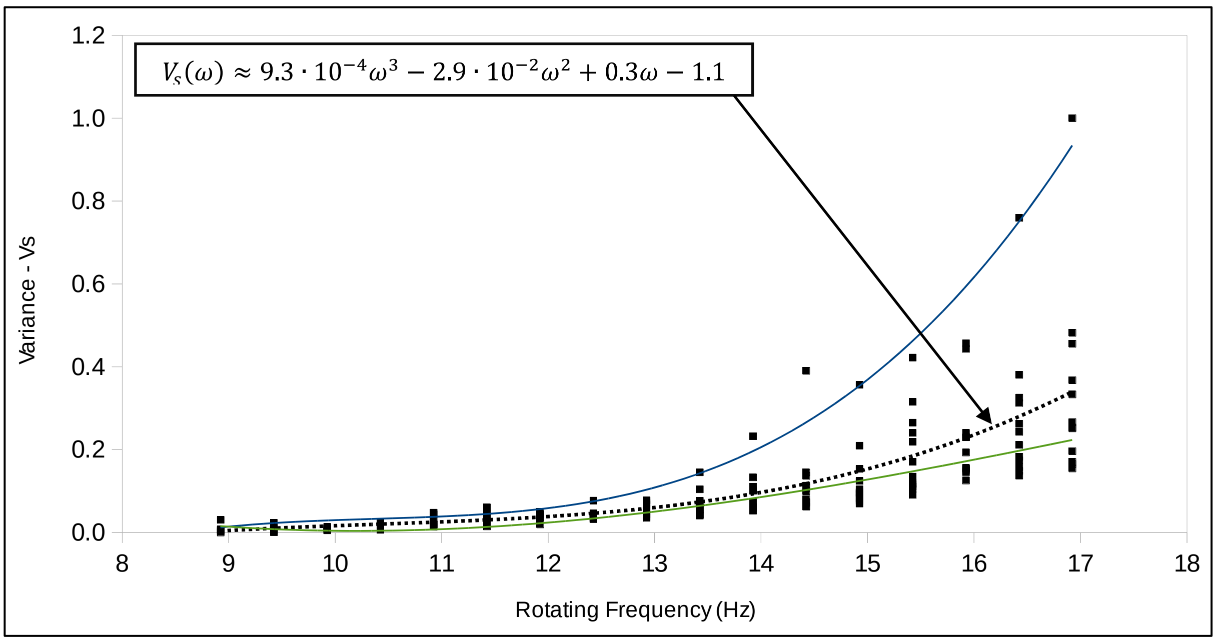


Figure 4.6: $\boldsymbol{V_s}$ as a function of $\boldsymbol{\omega}$. The blue and green trend lines correspond to the states with $\boldsymbol{\varphi_{\mathbf{air}} = 0}$ and $\boldsymbol{\varphi_{\mathbf{air}} \geq 5}$, respectively. The dotted black curve and best fit equation reflect the trend of all the states.

In this example, the correlation between V_s and can be suitably described by a third-degree polynomial. After determining the best fit function for the states, a refined variance $V_s' = V_s/V_s(\omega)$ can be calculated and renormalized.

Figure 4.7 shows the trend of V_s' with respect to φ_{air}.

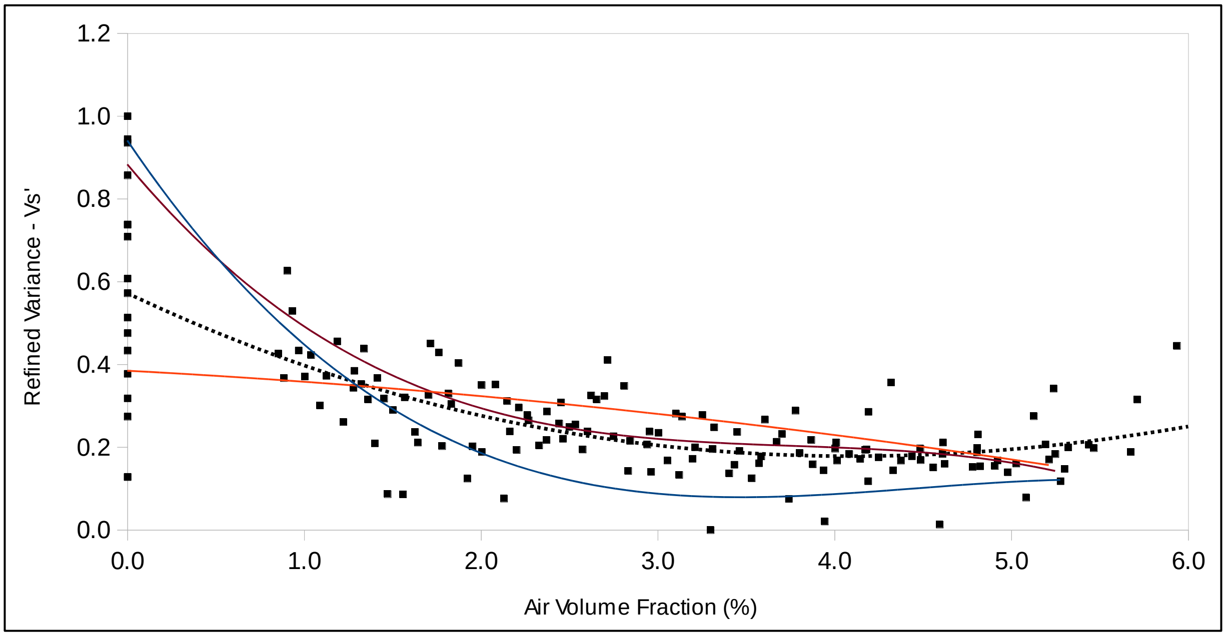


Figure 4.7: $V_s{}'$ as a function of φ_{air}. The blue, orange, and red trend lines correspond to $\omega = 8.9$, 12.9, and $16.9Hz$, respectively. The dashed black line reflects the trend of all the data points.

With the effects of ω reduced, it can be more clearly inferred that variance decreases universally with worsening gas entrainment. In contrast with the original feature, V_s' has increased variation across the experimental states and closer grouping about the overall trend (V_s exhibits $R^2 = 0.16$ about a third-order polynomial, whereas $V_s{}'$ has $R^2 = 0.40$).

This refinement process is applied to every input feature except for $\overline{p_{\text{out}}}$, which is retained specifically for its relationship to ω. The individual plots are shown in Appendix B2. Those that have correlations to both ω and φ_{air} are improved by varying degrees (e.g. V_s, E_R, and E_{FFT}). Those that have minimal correlation to ω are effectively unchanged (e.g. E_s, G_s, , and K_s). In the case where a feature appears poorly correlated to φ_{air} – namely, $\overline{p_{\text{out}}}$ – performing this calculation functions as a validation for that observation. Computing $\overline{p_{out}}(\omega)$ yields a near perfect linear correlation, and dividing that linear function out of $\overline{p_{\text{out}}}$ to produce $\overline{p_{\text{out}}}'$ reduces the feature to only noise. As such, it is omitted from the optimization.

With the refined set of input features, the binary, multi-class, and regression prediction models can be reevaluated. The results are shown in Table 4.6.

Table 4.6: Comparison of Gas Entrainment Classification Performance Before and After Feature Optimization

Prediction Model	Original Train/Test Accuracy (%)	Refined Train/Test Accuracy (%)
Binary MLP	90/84	94/90
Binary SVM	88/83	92/90
Multi-class MLP	82/59	85/62
Multi-class SVM	80/60	84/60
Regression	Half within 0.20/0.47	Half within 0.18/0.44

The impacts of the refined features are most prevalent in the binary prediction models. With the improved feature set, the binary MLP testing accuracy is improved to 90%, which meets the 90% threshold set in the research objectives. The binary SVM model improves to meet this threshold as well. The multi-class MLP model improves as well, placing 62% of the test states into the correct severity class. If its performance margin is relaxed to allow states that were misclassified by one class up or down, the testing accuracy rises to 88%. The multi-class SVM model exhibits no significant improvement in testing accuracy. Finally, the regression model demonstrates a minimal, but nonetheless positive change with the refined feature set. It predicts the gas entrainment severity within 0.44 of the true severity percentage in half of the testing states, and within 1.25 in 90%.

4.1.6 Error and Range of Application

While the prediction models achieve their intended purpose, there are additional methods that should be employed to improve their accuracy. The first proposed improvement is in how the data sets are assembled. In this study, the 310 training and testing states are aggregated manually, using a single dynamic pressure sample taken over one second at the given state. However, repeated measurements at the same operating condition reveal variability within each feature. This is shown in Figure 4.8.

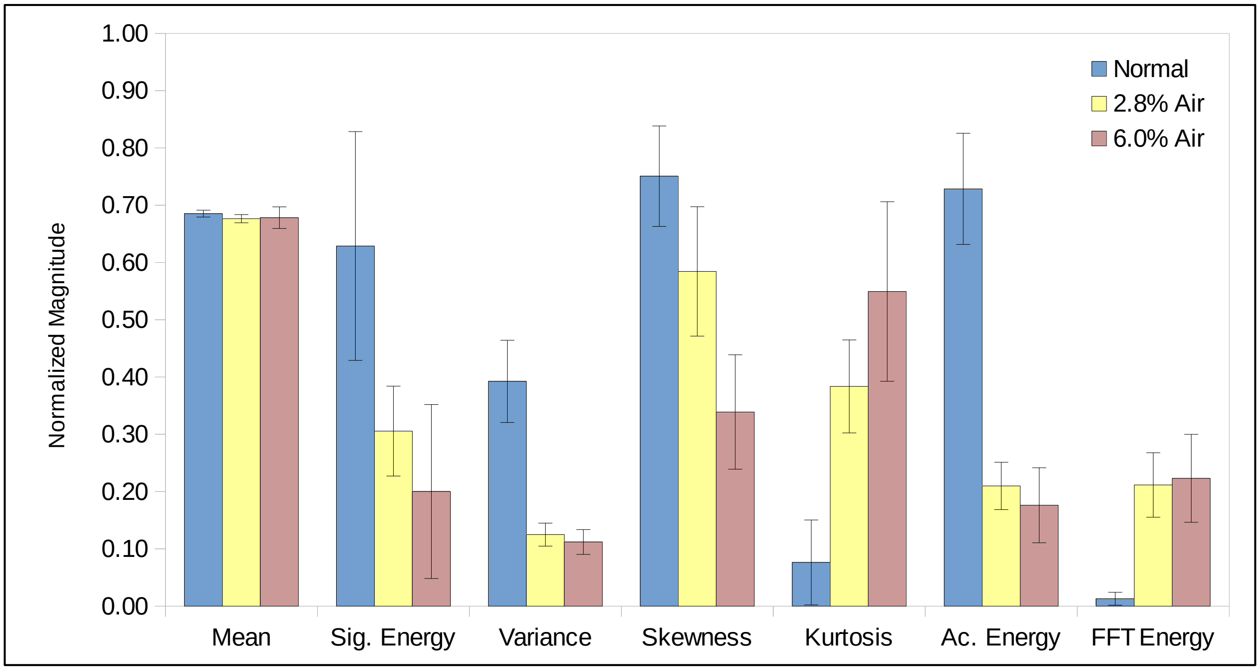


Figure 4.8: The average magnitudes of each input feature at $\boldsymbol{\omega = 14.9Hz}$, with $\boldsymbol{\varphi_{\mathrm{air}}}$ at 0% (blue bars), 2.8% (yellow bars) and 6.0% (red bars). These correspond to states 1M, 6M, and 12M from Table 4.1. Each magnitude is determined by averaging ten repeated 1 s pressure measurements of the same state. The error bars indicate an interval of two standard deviations.

This variability is due to the flow phenomena being somewhat inconsistent over the sample duration, rather than measurement error.[12] In the experimental data, the variability causes outliers. Though these outliers do not significantly skew the overall correlations between the features and gas entrainment, the classification algorithms can assign them undue weight, leading to erroneous predictions. To minimize this, the data should be composed of state measurements from significantly longer samples. This would have an averaging effect of the features in the training and testing sets, diminishing the influence of any transitory flow phenomena.

The second potential improvement is related to range of application. Across every model, states with lower impeller speeds are misclassified more frequently. The operating range is selected based on the viable ω values for the testbed system. However, in practice, a centrifugal pump operator is unlikely to use a centrifugal pump near its minimum output. It would be defensible to specify a narrower operating

[12] This can be inferred because each feature is calculated from the same measurement. If the variability were due primarily to errors in the transducer or data acquisition, each feature would exhibit similar variation.

band for the classification method, near the upper operating range of the pump, in which the gas entrainment predictions are more precise.

4.3 Impeller Wear Study

Radial impeller wear is simulated over 65 states, ranging from normal operation to $\alpha_{\text{imp}} = 3.0\%$. The training and testing sets contain 35 and 30 states, respectively. The values of ω and α_{imp} for each state are shown in Table 4.7, with the training states highlighted in gray. The states are indexed by row and column for reference.

Table 4.7: Simulated Radial Impeller Wear States

		A	*B*	*C*	*D*	*E*
	ω (Hz) →	**8.75**	**13.75**	**18.75**	**23.75**	**28.75**
1		0.0	0.0	0.0	0.0	0.0
2		0.2	0.2	0.2	0.2	0.2
3		0.5	0.5	0.5	0.5	0.5
4		0.7	0.7	0.7	0.7	0.7
5		1.0	1.0	1.0	1.0	1.0
6		1.2	1.2	1.2	1.2	1.2
7	**α_{imp} (%) →**	1.5	1.5	1.5	1.5	1.5
8		1.7	1.7	1.7	1.7	1.7
9		2.0	2.0	2.0	2.0	2.0
10		2.2	2.2	2.2	2.2	2.2
11		2.5	2.5	2.5	2.5	2.5
12		2.7	2.7	2.7	2.7	2.7
13		3.0	3.0	3.0	3.0	3.0

Figure 4.9 shows the simulated waveforms of p_{out} at $\alpha_{\text{imp}} = 0$ and $\alpha_{\text{imp}} = 3.0\%$ (rows 1 and 13 from Table 4.7).

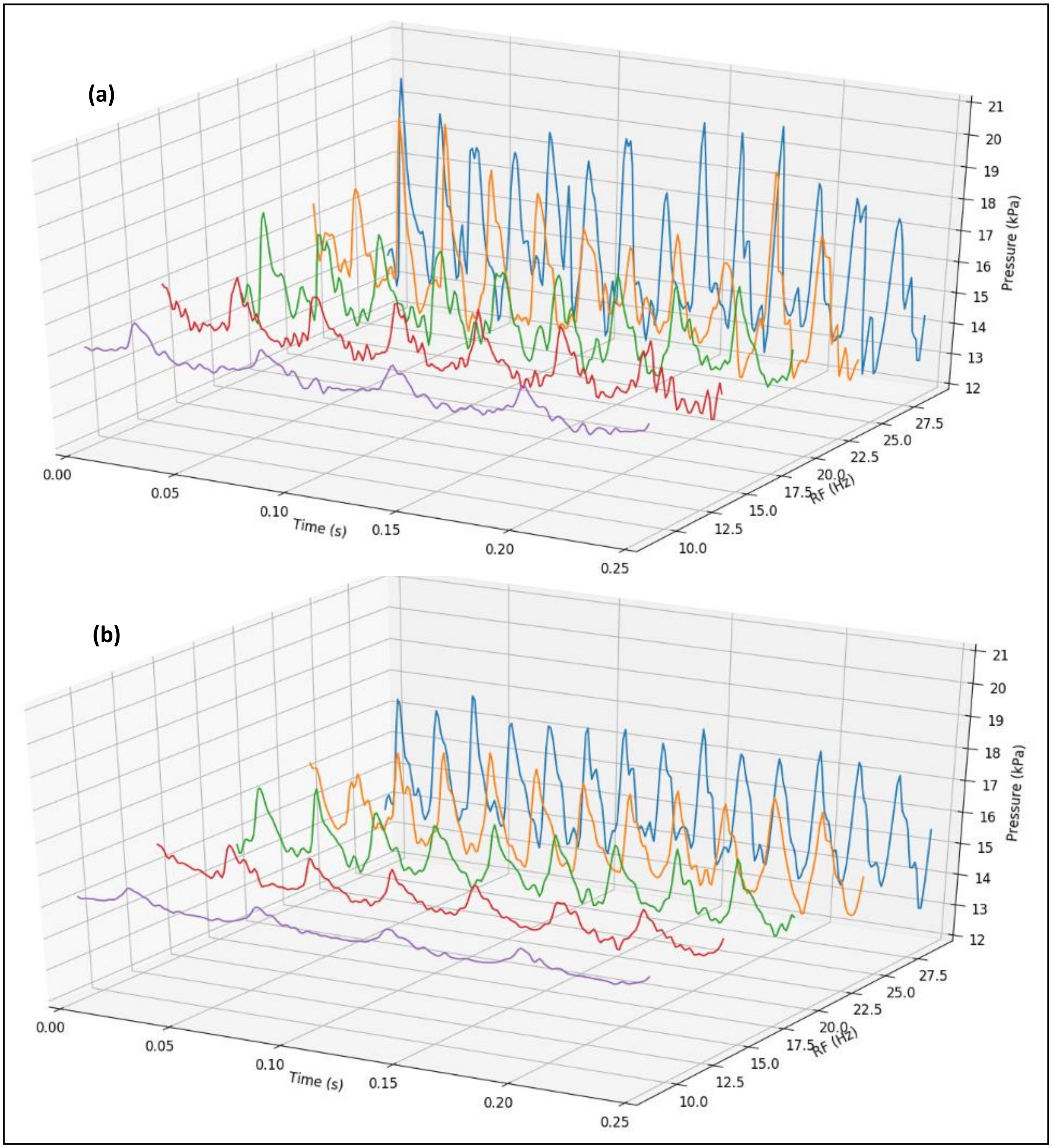


Figure 4.9: Simulated discharge pressure fluctuations across and range of rotating frequencies during a) normal operation and b) operation with $\boldsymbol{\alpha}_{\mathbf{imp}} = \mathbf{3.0}\%$.

The simulated pressure waveforms in Figure 4.9a mirror discharge pressure fluctuations reported by Cao [59], S. Chu *et al.* [74], and Higham [47]. In the normal operating condition, at $\omega = 8.75Hz$, the main mode of the discharge pressure fluctuation varies within a band of 1 kPa.[13] When ω is increased to 28.75 Hz, the signal variation grows to approximately 7 kPa. Each sample in Figure 4.9a exhibits a degree of high frequency fluctuation, which is interpreted as a combined effect of fluid turbulence and numerical noise generated by the necessarily coarse mesh.

In Figure 4.9b, the effects of radial erosion of the impeller manifest primarily as reduced variance in p_{out}. Unlike gas entrainment, impeller wear does not appear to dramatically alter the shape of the pressure pulses, other than reducing their amplitude. However, at the highest frequencies (orange and blue trends), it is observed that the pressure peaks have increasing variability from peak to peak, which suggests that low frequency fluctuations may become more prevalent.

4.3.1 Feature Correlation

As with the gas entrainment study, the signal decomposition methods from Section 3.2 are employed to reduce the simulated measurements of p_{out} into the desired feature space. The resulting feature trends are shown in Figures 4.10a through 4.10f. The trend for $\overline{p_{\mathrm{out}}}$ is excluded, as it is a function of the simulation's boundary condition and not an external phenomenon. Each figure displays four second-order polynomial best-fit lines corresponding to the features associated with i) the lowest impeller speed, $\omega = 8.75Hz$ (blue line), ii) the median speed, $\omega = 18.75Hz$ (orange line), iii) the highest speed, $\omega = 28.75Hz$ (maroon line), and iv) all the data points (black dashed line). These are intended to illustrate cases where the feature has a dual dependence on impeller wear and ω.

[13] Due the considerable geometric differences between the numerical model and physical testbed centrifugal pump, the raw pressure magnitudes between the entrainment and wear studies should be interpreted independently.

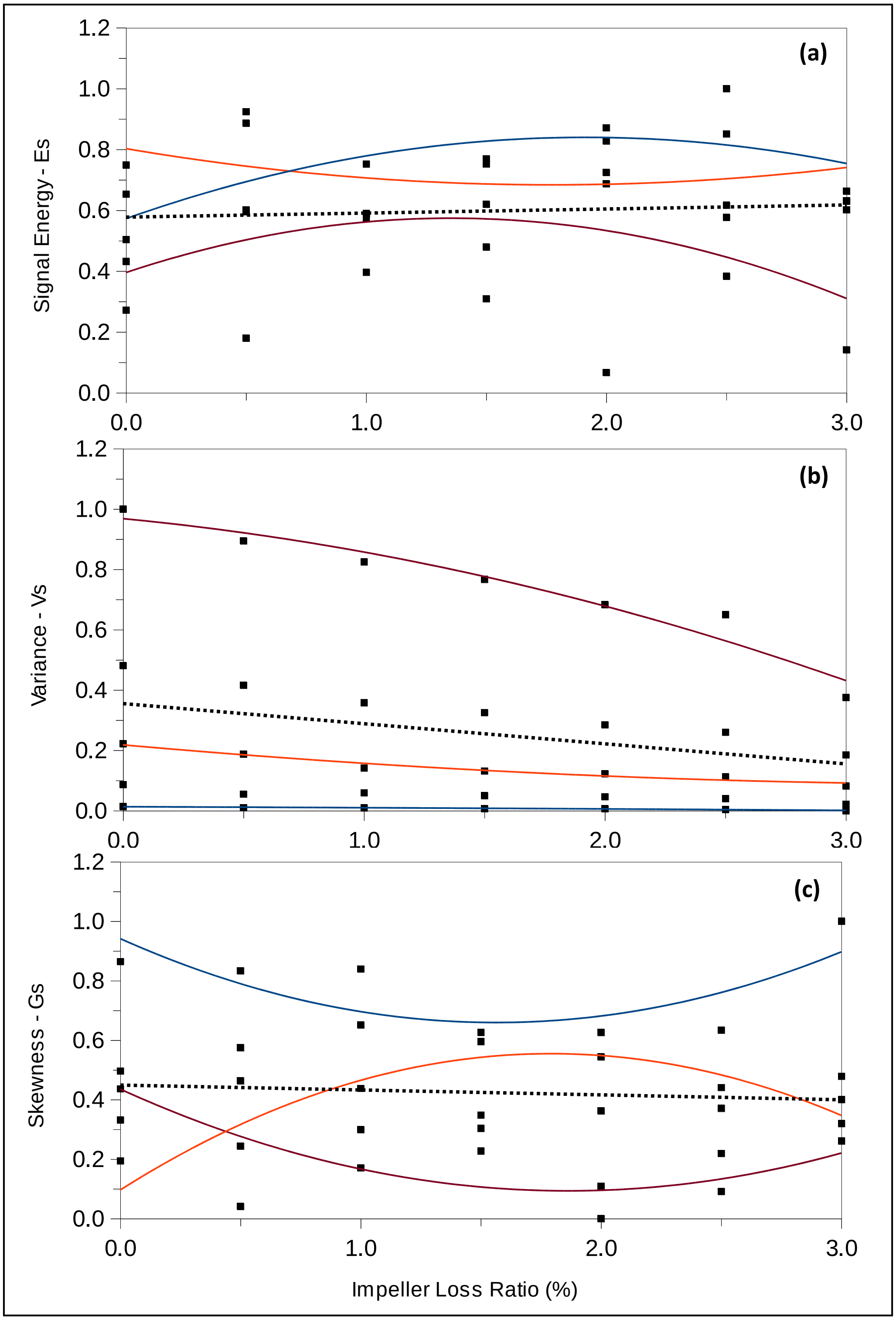


Figure 4.10a-c: The normalized values of a) E_s, b) V_s, and c) G_s as a function of α_{imp}.

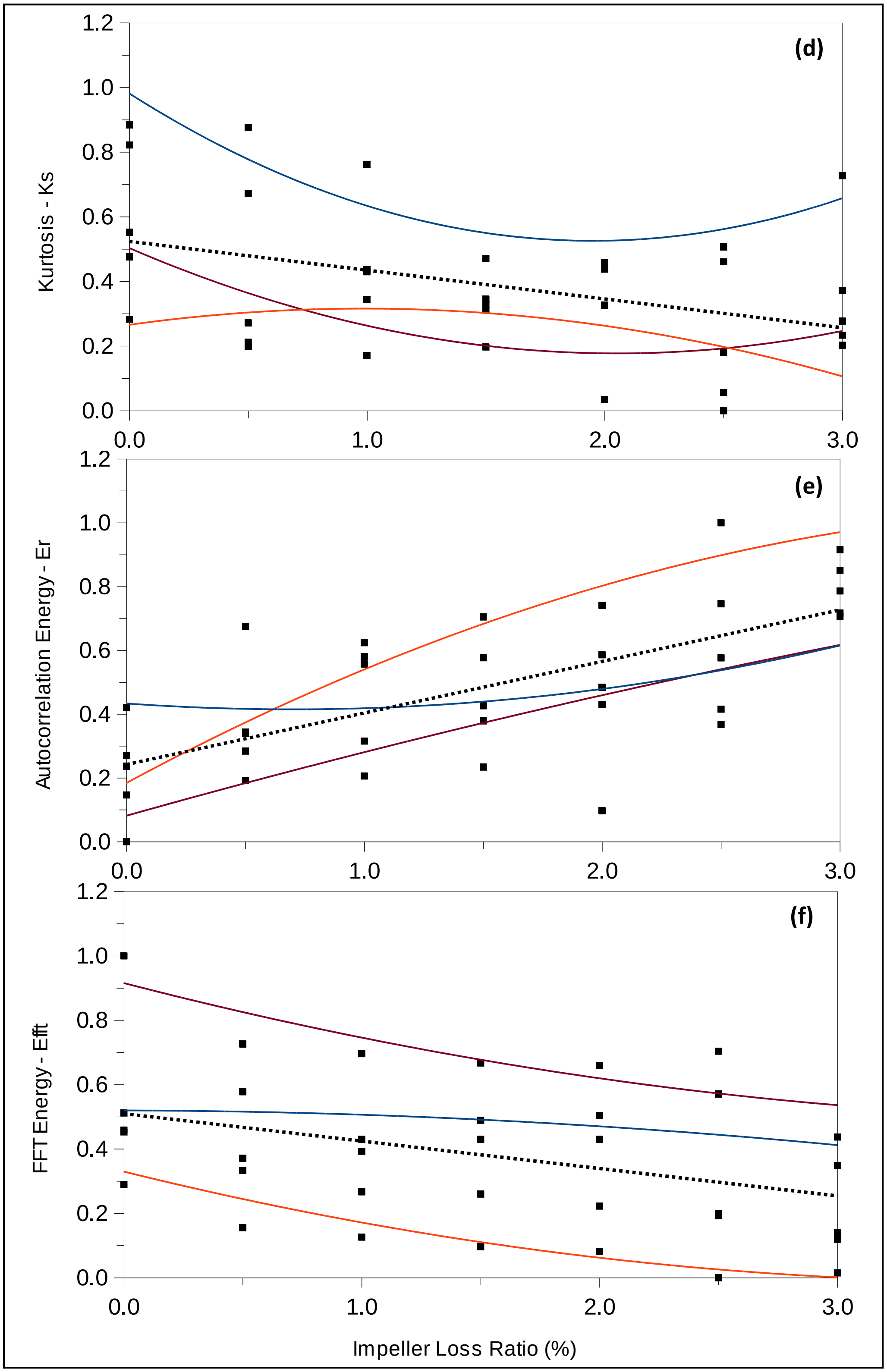
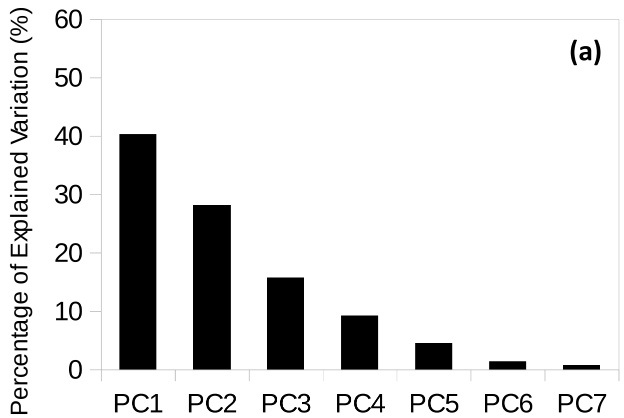


Figure 4.10d-f: The normalized values of d) $\boldsymbol{K_s}$, e) $\boldsymbol{E_R}$, and f) $\boldsymbol{E_{\mathrm{FFT}}}$ as a function of $\boldsymbol{\alpha_{\mathrm{imp}}}$.

Figure 4.10 suggests that some of the features within the training data are less strongly correlated to radial impeller wear than gas entrainment. Signal energy E_{S} and skewness G_{S} (Figures 4.10a and 4.10c), for example, demonstrate no coherent relationship overall, which is consistent with the results from the preliminary simulations. Kurtosis K_{S} (Figure 4.10d) appears to follow a slightly negative trend as α_{imp} increases, but the correlation is marginal. The training set is not extensive enough to conclusively state that these three features are universally uncorrelated to the target condition, but in the context of the presented feature space, it is unlikely that they will contribute significantly to classification.

In Figure 4.10b, V_{S} follows a consistent, decreasing trend. The additional correlation to impeller speed is also apparent, causing a vertical spread between the slow, intermediate, and high speed states. This infers that the feature optimization described in Section 4.1.5 may be advantageous. Autocorrelation energy E_{R} (Figure 4.10e) and FFT energy E_{FFT} (Figure 4.10f) also exhibit consistent variation with respect to α_{imp} and may benefit from optimization as well.

The input features are further evaluated using PCA. The resulting scree plot is shown in Figure 4.11a, along with the loadings for PC1 and PC2. A comparison of PC1 and PC2 is shown in Figure 4.11b.



	Loading (PC1)	Loading (PC2)
E_s	0.534	0.015
E_R	0.433	0.436
E_FFT	0.423	0.421
G_s	0.413	0.398
V_s	0.329	0.133
K_s	0.200	0.651
$\overline{p}_\mathrm{out}$	0.175	0.178

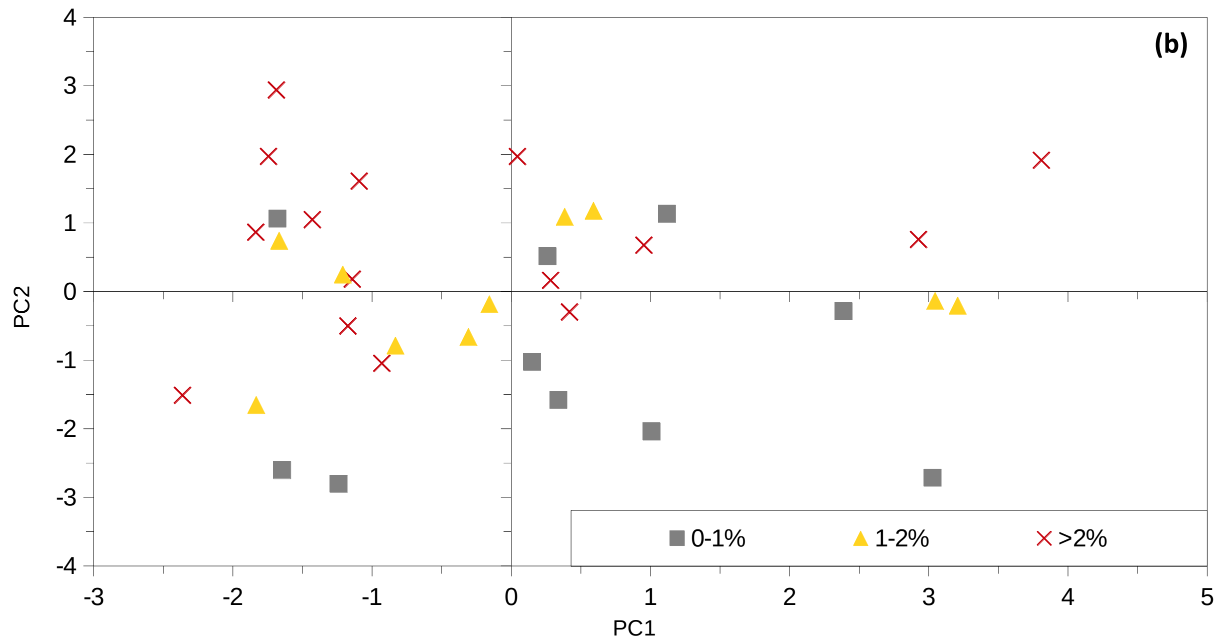


Figure 4.11: a) A scree plot of the PCs of radial impeller wear and the loadings for each variable with respect to PC1 and PC2, listed in order of their contribution to PC1; b) The distribution of the experimental states across the first two PCs.

Figure 4.11a shows that approximately 70% of the explained variation across the features is accounted for by PC1 and PC2. The PC1 loadings show E_s, E_R, and E_FFT, as the dominant contributors. The scatterplot in Figure 4.11b establishes that the varying states of impeller wear are not linearly separable based on the feature space.

4.3.2 Binary Classification

With the acceptability threshold $\tau_\mathrm{wear} = 1.5\%$, the training and testing accuracies of the binary MLP are 93% and 87%, respectively. This is achieved using a learning rate of 0.02 and $7{\cdot}10^4$ training iterations. The binary SVM model achieves equivalent training and testing accuracies. The resulting truth table for the binary MLP testing data is shown in Table 4.8.

Table 4.8: Truth Table for Binary MLP Classification of Impeller Wear (Testing)

True Healthy: **18**	False Healthy: **2**
False Degraded: **2**	True Degraded: **8**

Of the thirty training states, four are misclassified, with an even split between false healthy and false degraded predictions. Each incorrect classification occurs near τ_{wear}, as shown (underlined) in Table 4.9.

Table 4.9: Impeller Wear States Misclassified by Binary MLP Model (Testing)

		A	*B*	*C*	*D*	*E*
	$\boldsymbol{\omega}$ (Hz) →	8.75	13.75	18.75	23.75	28.75
1		0.2	0.2	0.2	0.2	0.2
2		0.7	0.7	0.7	0.7	0.7
3	$\boldsymbol{\alpha_{\mathrm{imp}}}$ (%) →	1.2	1.2	1.2	1.2	1.2
4		1.7	$\underline{1.7}$	1.7	$\underline{1.7}$	1.7
5		2.2	$\underline{2.2}$	2.2	$\underline{2.2}$	2.2
6		2.7	2.7	2.7	2.7	2.7

The binary MLP and SVM models nearly achieve the 90% accuracy goal in the stated research objectives, despite minimal contribution from three of the six input features. However, there is still margin for improvement by reducing the influence of ω on the input features.

4.3.2 Multi-class Classification

The multi-class MLP model places each training state in the correct group with 89% accuracy. The corresponding accuracy in the testing data is 70%. This discrepancy infers that the model may be on the threshold of overfitting the training data. Testing error is exacerbated with more training iterations. It is suspected that this is a consequence of the relatively small pool of points within each class in the training data set. The multi-class SVM model achieves training/testing prediction accuracies of 87% and 73%, respectively.

The misclassified testing states from the multi-class MLP model are shown in Table 4.10.

Table 4.10: Impeller Wear States Misclassified by Multi-class MLP Model (Testing)

		A	B	C	D	E
	$\boldsymbol{\omega}$ (Hz) →	8.75	13.75	18.75	23.75	28.75
1	α_{imp} (%) →	0.2	0.2	0.2	0.2	0.2
2		0.7	0.7	0.7	0.7	0.7
3		1.2	1.2	1.2	1.2	1.2
4		1.7	1.7	1.7	1.7	1.7
5		2.2	2.2	2.2	2.2	2.2
6		2.7	2.7	2.7	2.7	2.7

Overfitting can likely be reduced using a larger training set and refining the input features to improve their correlation to the target condition.

4.3.3 Regression

The regression model performs moderately well, with a median training prediction error of 0.22 (standard deviation 0.20) of the true percentage value of α_{imp} and a median testing prediction error of 0.30 (standard deviation 0.31). The testing predictions with error greater than 0.6 from the true α_{imp} are shown in Table 4.11.

Table 4.11: Impeller Wear States Severely Misclassified by Regression Model (Testing)

		A	B	C	D	E
	$\boldsymbol{\omega}$ (Hz) →	8.75	13.75	18.75	23.75	28.75
2	α_{imp} (%) →	0.2	0.2	0.2	0.2	0.2
4		0.7	0.7	0.7	0.7	0.7
6		1.2	1.2	1.2	1.2	1.2
8		1.7	1.7	1.7	1.7	1.7
10		2.2	2.2	2.2	2.2	2.2
12		2.7	2.7	2.7	2.7	2.7

4.3.4 Feature Optimization

Using the method described in Section 4.1.5, each input feature is refined to reduce its dependence on ω. The impact is most significant on V_s, shown in Figure 4.12.

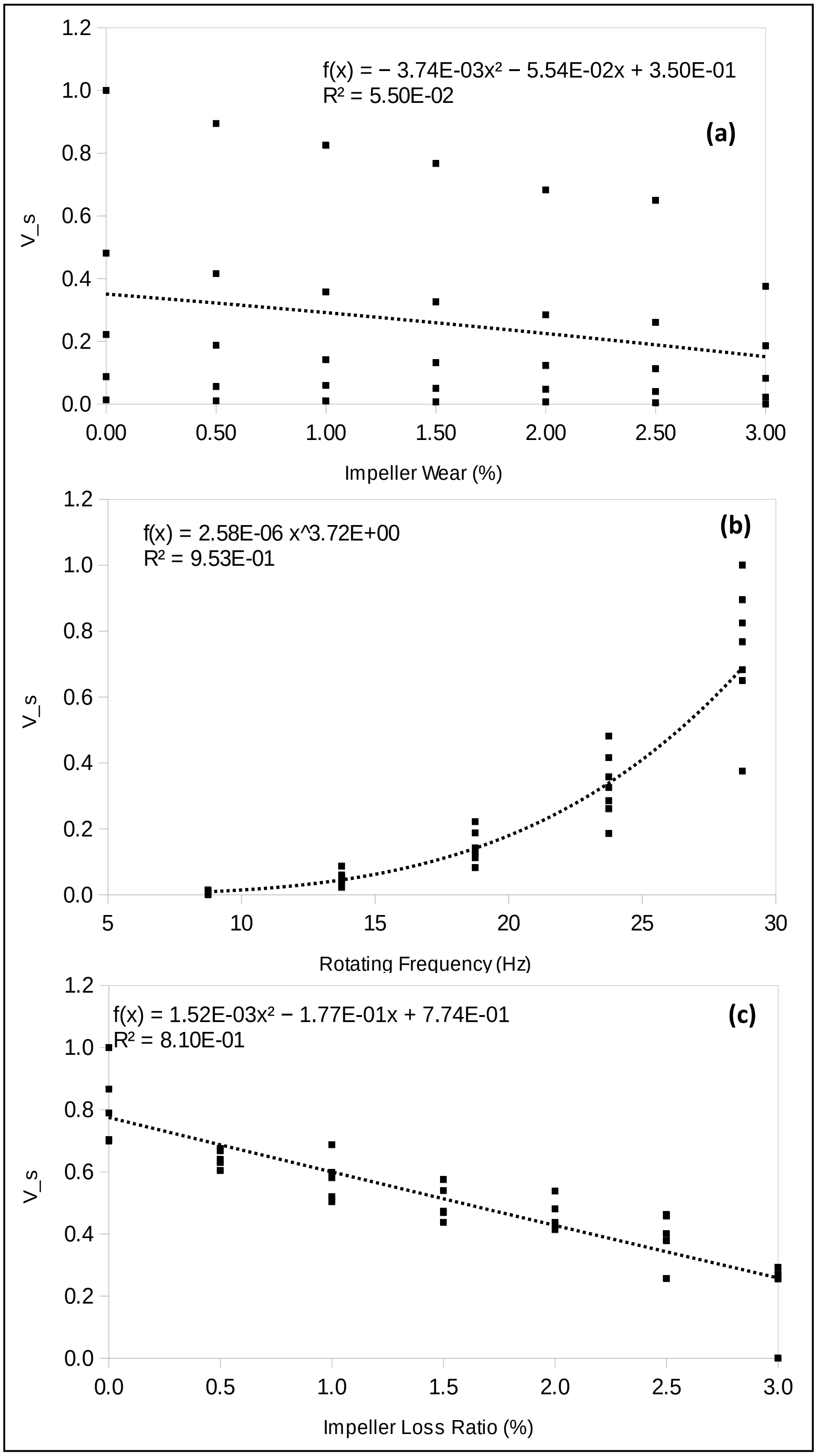


Figure 4.12: The optimization steps of $\boldsymbol{V_s}$: a) $\boldsymbol{V_s}$ as a function of $\boldsymbol{\alpha_{imp}}$; b) $\boldsymbol{V_s}$ as a function of $\boldsymbol{\omega}$; c) $\boldsymbol{V_s}'$ as a function of $\boldsymbol{\alpha_{imp}}$.

The remaining feature plots are shown in Appendix B2. The features $\overline{p_{\text{out}}}$, E_{s}, G_{s}, and K_{s} are largely independent of ω and exhibit minimal change. However, the correlations of V_{s}, E_{R}, and E_{FFT} are strongly enhanced. This impact is demonstrated when the new feature set is used for classification, as shown in Table 4.12.

Table 4.12: Comparison of Impeller Wear Classification Performance Before and After Feature Optimization

Prediction Model	Original Train/Test Accuracy (%)	Refined Train/Test Accuracy (%)
Binary MLP	93/87	100/97
Binary SVM	93/87	97/93
Multi-class MLP	87/70	92/82
Multi-class SVM	89/73	89/82
Regression	Half within 0.22/0.30	Half within 0.11/0.16

With the optimized feature set, all classification methods have better accuracy. For binary classification with the MLP, the training and testing accuracy jump to 100% and 97%, respectively. The training/testing accuracy of the multi-class MLP improve to 92/82%. Notably, the median error of the regression model halves for both training and testing data, classifying half of the points in the testing set within 0.16 of their true wear extent. These dramatic performance improvements reinforce the importance of employing input features with robust correlation to the target phenomenon.

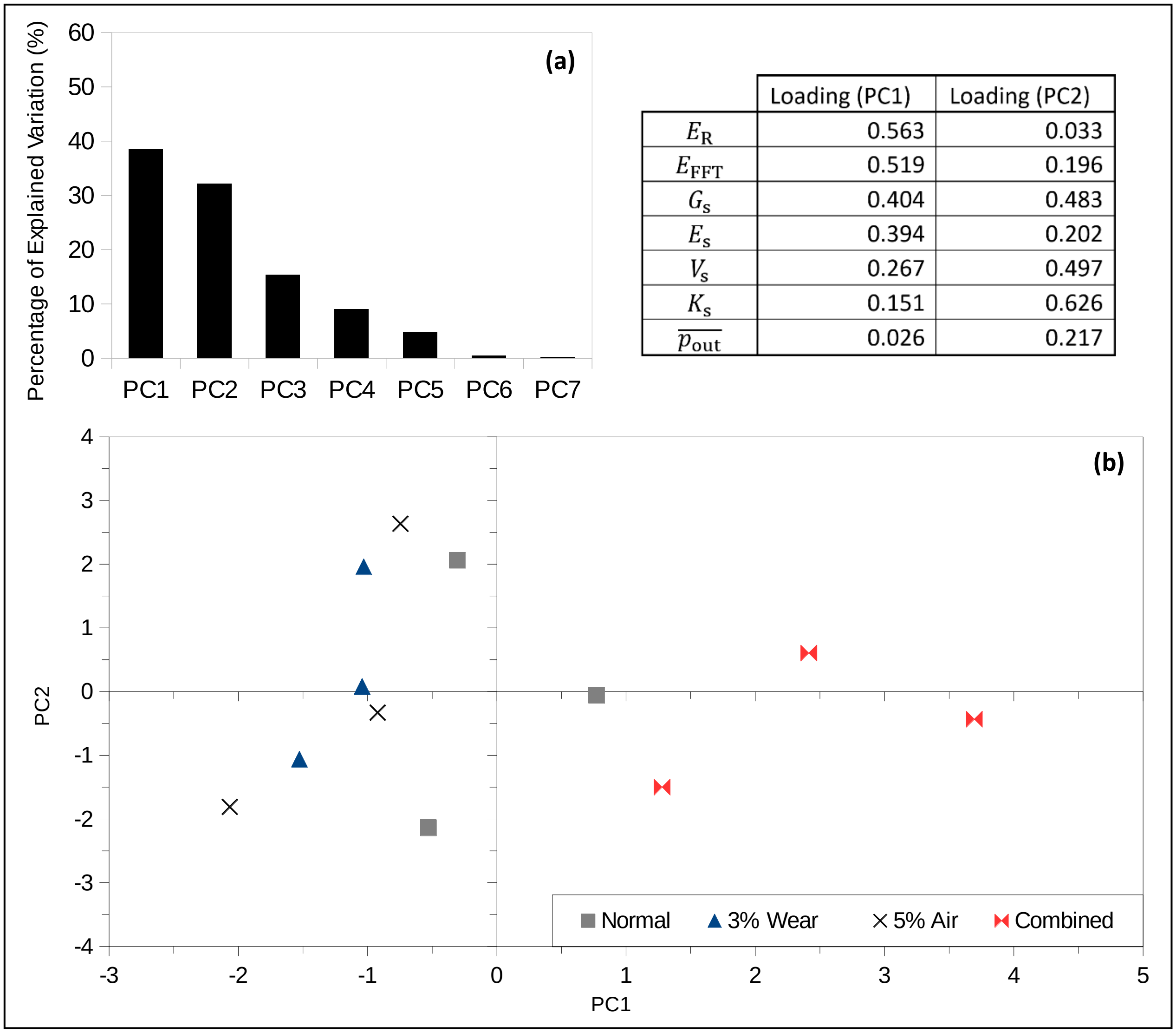
4.4 Combined Study

A total of twelve states are simulated in the combined study. The set comprises states from normal operation, severe impeller wear ($\alpha_{\text{imp}} = 3.0\%$), severe gas entrainment ($\varphi_{\text{air}} = 5.0\%$), and the simultaneous occurrence of both. They are shown in Table 4.13.

Table 4.13: Simulated States for Combined Study

		A	*B*	*C*
	$\boldsymbol{\omega}$ (Hz) →	8.75	13.75	28.75
1	Normal	-	-	-
2	$\boldsymbol{\alpha_{\text{imp}}}$ (%) →	3.0	3.0	3.0
3	$\boldsymbol{\varphi_{\text{air}}}$ (%) →	5.0	5.0	5.0
4	Combined $\boldsymbol{\alpha_{\text{imp}}/\varphi_{\text{air}}}$ (%)	3.0/5.0	3.0/5.0	3.0/5.0

The small pool of states precludes the application of machine learning to characterize the conditions. PCA is applied to evaluate the correlation of the input features and the linear separability of the states. The resulting scree plot and loadings for PC1 and PC2 are shown in Figure 4.13a. The states are plotted across the first two PCs in Figure 4.13b.



	Loading (PC1)	Loading (PC2)
E_{R}	0.563	0.033
E_{FFT}	0.519	0.196
G_{s}	0.404	0.483
E_{s}	0.394	0.202
V_{s}	0.267	0.497
K_{s}	0.151	0.626
$\overline{p_{\mathrm{out}}}$	0.026	0.217

Figure 4.13: a) A scree plot of the PCs of normal operation, severe impeller wear, severe air entrainment, and combined conditions. The loading for each variable with respect to PC1 is adjacent, ranked in descending order; b) The distribution of the various conditions across the first two PCs.

Figure 4.13a shows that PC1 and PC2 represent approximately 70% of the explained variation between the states. The loadings for PC1 suggest that E_{R}, E_{FFT}, and G_{s} are the dominant contributors.

Projecting the states onto PC1 and PC2 reveals relatively poor separation between normal, gas entrainment, and impeller wear states. This follows the results from the two previous studies, which suggested that the individual conditions are not linearly separable. However, when the adverse conditions manifest simultaneously, the states fall in a distinct region (to the right of PC1). This supports that the input features are satisfactory for distinguishing individual and combined phenomena in their severe states. However, a significantly larger data set is required to validate their classification performance on intermediate combined-phenomena states.

5. CONCLUSIONS AND FUTURE WORK

This chapter recapitulates the results presented in this thesis, applications and limitations of the diagnostic method, the original research objectives, conclusions, and recommendations for future research related to centrifugal pump monitoring via dynamic pressure measurements.

5.1 Summary of Results

This work describes a novel method for characterizing the severity of adverse operating conditions in a centrifugal pump using dynamic pressure measurements collected at the discharge. The method is validated in respective studies against two pervasive phenomena; gas entrainment and radial erosion of the impeller blades. First, simulations of the two target phenomena are used to generate representative pressure measurements. These are, in turn, employed to propose a unique set of characteristic statistical features, with correlations to both phenomena. The features include signal mean, energy, variance, skewness, kurtosis, energy of the autocorrelation function, and total spectral energy. Based on the respective trends from the preliminary simulations, three machine learning methods are put forth, each with increasing resolution of the target phenomenon. The methods are; binary classification using a multi-layer perceptron (MLP), multi-class classification using a MLP, and continuous-value prediction using a random forest regression algorithm.

The method is validated in two principal studies. In the first, experimental pressure measurements are collected in 310 states of varying impeller speed and gas entrainment severity, up to 6% void fraction of air. The states are divided in to two equal sets of 155 states for training and testing of the prediction algorithms. Using the base feature set, the binary MLP successfully predicts gas entrainment exceeding 2% void fraction in the testing set with 84% accuracy. The multi-class MLP places the testing states into the correct 1% severity band with 59% accuracy, with 70% of the misclassifications falling just one severity band up or down. Binary and multi-class support vector machines (SVM) models, implemented in parallel for validation, yield comparable prediction accuracies. The random forest regression model has a median error of 0.47 of the true gas entrainment severity percentage.

To improve the prediction performance, an optimization method is proposed that refines the correlations between the input features and gas entrainment by reducing their dual dependence on impeller speed. In doing so, the testing performance of the binary and multi-class models is improved to

90% and 62%, respectively. The median error of the regression model reduces to within 0.44 of the true entrainment severity.

In the second study, pressure measurements are simulated for 65 states of varying impeller speed and radial wear, up to 3% diametric loss. The states are divided 35/30 to form the training and testing sets. The binary MLP successfully predicts impeller wear exceeding 1.5% in 87% of the testing states, which improves to 97% after optimizing the input features. The multi-class MLP places the testing samples into the correct 1% severity band with 70% accuracy, improving to 82% after optimizing. The median prediction error of the regression model before and after refinement is 0.30 and 0.16 of the true wear severity percentage, respectively.

In addition, a preliminary study is presented concerning the ability of the diagnostic method to differentiate between simultaneously occurring gas entrainment and impeller wear. The initial simulations advocate its feasibility, and a follow-up experiment is proposed.

5.2 Applications and Limitations

As discussed in the introduction, improving centrifugal pump performance monitoring is as much a problem of practicability as outright technology. Cost, adaptability, scalability, and long term accuracy are as fundamentally important as the sensing method by which performance is measured. Existing commercial solutions address this balance poorly, yet the development of elaborate pump diagnostic tools which are impractical for wide scale use is perpetual. The void of practical alternatives has led to over reliance on rudimentary, static performance references, such as pump performance curve charts, which can lose accuracy with process changes. The end consequence is that many centrifugal pumps around the world operate in unknown, often inefficient states.

Condition monitoring via dynamic pressure measurements and machine learning presents an opportunity to improve the state of the art in both accessibility and versatility. A principal advantage of the method discussed in this research is that the process would fall primarily to the pump manufacturer to employ, rather than the individual user. The classification models would be trained and configured by the manufacturer, then provided as a supplementary diagnostic service for the pump operator. The operator need only invest in a single transducer per pump, make dynamic measurements following the manufacturer's method, and observe the classification results.

Here, gas entrainment and radial impeller wear are studied because of their relative difficulty to characterize with existing systems. However, in practice, there could be many target conditions to be classified using this method. The main constraint would be in the variety of adverse conditions the pump manufacturer could generate to amass reliable training data.

The classification technique is not without limitations, however. The studies in this thesis satisfactorily demonstrate classification of adverse pump phenomena via dynamic pressure sensing as a concept, but certain industrial processes may require resolution beyond what has been reported. Achieving this classification accuracy would require substantially larger sets of longer state measurements, and potentially additional refinement and expansion of the feature set. A second limitation is that, as a fluid-based sensing approach, this method is not well suited for classifying the many mechanical faults that can impact the performance of centrifugal pumps. This could be remedied by employing a combined pressure-sensor/accelerometer and employing features based on structural vibrations.

5.3 Conclusions

The original objectives for this research are presented in Section 1.6. Based on the presented studies, we conclude the following:

(1) Statistical Features: We have proposed a set of statistical measures that can be used to quantify the target phenomena based solely on dynamic discharge pressure measurements.

(2) Correlation: The relevance of those measures, as associated to detecting the target phenomena, has been evaluated. Further, that evaluation has also been leveraged to improve the correlation of the statistical features to the target phenomena.

(3) State Prediction: Using a binary MLP, experimental gas-entrainment states exceeding a 2% volume fraction of air have been classified with 90% accuracy across varying impeller speeds. Simulated radial wear exceeding 1.5% of the impeller diameter has been classified with 97% accuracy using the same algorithm.

(4) Refined Characterization: We have demonstrated satisfactory prediction using multi-class classification via a MLP and continuous-value regression via random forest algorithm.

(5) Accessibility: The developed diagnostic method employs a solitary, conventional pressure transducer that costs $115 and has a rise time of 2 ms.

(6) Implementation: Approaches for industrial implementation have been discussed.

This research has established that dynamic measurement of the discharge pressure fluctuations from a centrifugal pump is a viable method to evaluate fluid phenomena. We have shown that the severity of the target phenomena can be categorized using a set of characteristic statistical measures derived from the discharge pressure fluctuations. Reducing these features' dependence on extraneous input variables (specifically, impeller speed), improves their correlation to the target phenomena. It has also been established that prediction performance is reliant on a suitably extensive and diverse set of training conditions.

Finally, we have demonstrated the method using an affordable, universally accessible sensor.

5.5 Future Work

Although this work has demonstrated the validity of characterizing adverse centrifugal pump phenomena using dynamic discharge pressure measurements, there are a variety of avenues by which the research could be augmented. First, it is recommended that trials be conducted to validate the impeller wear classification method using experimental measurements. Additionally, it would be beneficial to automate the experimental setup in order to expediently recreate the experiments using a considerably larger set of training states (i.e. thousands, rather than hundreds or tens), using longer samples of the target conditions. This would allow for a more robust evaluation of the input features, and could potentially yield better optimization of the feature correlations or additional statistical measures beyond the set discussed herein. Such an automated setup would also allow for a thorough investigation of combined-phenomena states, which are preliminarily discussed here.

Our research focuses specifically on gas entrainment and impeller wear, but there is a potential to expand the approach to include a variety of other fluid phenomena, adverse or otherwise. It would be valuable to investigate the full spectrum of fluid behaviors that can be characterized using this technique, and explore the statistical means necessary to observe and characterize their correlations.

BIBLIOGRAPHY

[1] M. L. Adams, *Power plant centrifugal pumps: problem analysis and troubleshooting*, 2nd edition. Boca Raton: Taylor & Francis, a CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa, plc, 2016.

[2] J. Pemberton and J. Bachmann, "Pump System Performance Impacts," *Eng. Min. J.*, vol. April 2010, no. 211, pp. 56–59, 2010.

[3] European Commission, "Study on Improving the Energy Efficiency of Pumps," European Commission, 2001.

[4] P. G. Kini, R. C. Bansal, and R. S. Aithal, "Performance Analysis of Centrifugal Pumps Subjected to Voltage Variation and Unbalance," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 562–569, 2008, doi: 10.1109/TIE.2007.911947.

[5] M. Šavar, H. Kozmar, and I. Sutlović, "Improving centrifugal pump efficiency by impeller trimming," *Desalination*, vol. 249, no. 2, pp. 654–659, Dec. 2009, doi: 10.1016/j.desal.2008.11.018.

[6] S. Krüger, N. Martin, and P. Dupont, "Assessment of Wear Erosion in Pump Impellers," *Proc. Twenty-Sixth Int. Pump Users Symp.*, pp. 51–56, 2010.

[7] ERMP, "Performance Monitoring Discussion with BC Pump Manufacturer," Nov. 2017.

[8] J. Zhang, H. Fan, W. Zhang, and Z. Xie, "Energy performance and flow characteristics of a multiphase pump with different tip clearance sizes," *Adv. Mech. Eng.*, vol. 11, no. 1, p. 168781401882335, Jan. 2019, doi: 10.1177/1687814018823356.

[9] J. F. Gülich, *Centrifugal pumps*. Berlin ; New York: Springer, 2008.

[10] S. X. Chen, Z. Y. Pan, Y. L. Wu, and D. Q. Zhang, "Simulation and experiment of the effect of clearance of impeller wear-rings on the performance of centrifugal pump," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 15, no. 7, p. 072017, Nov. 2012, doi: 10.1088/1755-1315/15/7/072017.

[11] The Gorman-Rupp Company, "VG1 Pump Specification." The Gorman-Rupp Company, 2016.

[12] H. P. Bloch, *Pump wisdom: problem solving for operators and specialists*. Hoboken, N.J: Wiley, 2011.

[13] ERMP, "Energy Reduction Discussion with BC Paper Producer," Nov. 2017.

[14] International Organization for Standardization, *ISO 9906:2012: Rotodynamic pumps — Hydraulic performance acceptance tests — Grades 1, 2 and 3*. International Organization for Standardization, 2012.

[15] R. Tarodiya and B. K. Gandhi, "Hydraulic performance and erosive wear of centrifugal slurry pumps - A review," *Powder Technol.*, vol. 305, pp. 27–38, Jan. 2017, doi: 10.1016/j.powtec.2016.09.048.

[16] Thermancy, "Acuity Wireless Sensor System," 2017, [Online]. Available: http://www.thermancy.co.uk/products/products-fixed-wireless/.

[17] Robertson, "P22 Portable Pump Monitor," 2018, [Online]. Available: https://www.robertson.technology/p22-portable-pump-monitor.

[18] A. E. Cattaert, “High pressure pump efficiency determination from temperature and pressure measurements,” in *2007 IEEE Power Engineering Society Conference and Exposition in Africa-PowerAfrica*, 2007, pp. 1–8.

[19] D. Shang, “Application Research on Testing Efficiency of Main Drainage Pump in Coal Mine Using Thermodynamic Theories,” *Int. J. Rotating Mach.*, vol. 2017, pp. 1–8, 2017, doi: 10.1155/2017/5936506.

[20] Austin. Whillier, “Pump Efficiency Determination in Chemical Plant from Temperature Measurements,” *Ind. Eng. Chem. Process Des. Dev.*, vol. 7, no. 2, pp. 194–196, Apr. 1968, doi: 10.1021/i260026a007.

[21] British Standards Institution, *ISO 5198:1999: Centrifugal, mixed flow and axial pumps: code for hydraulic performance tests : precision class*. London: British Standards Institution, 1999.

[22] A. Milne, “Considerations when using the Thermodynamic Method of Pump Testing,” *World Pumps*, no. Special Feature, Jul. 1990.

[23] D. F. Niculescu, A. Olaru, A. Ghionea, M. Hajduk, and M. Vagaš, “Assisted Research of the Centrifugal Pump’s Vibration Behavior,” *Appl. Mech. Mater.*, vol. 436, pp. 85–91, Oct. 2013, doi: 10.4028/www.scientific.net/AMM.436.85.

[24] Y. Luo, Y. Han, S. Yuan, and J. Yuan, “Research on the Single-Value Indicators for Centrifugal Pump Based on Vibration Signals,” p. 18, 2020.

[25] K. Wang, X. Lu, and X. He, “Experimental Investigation of Vibration Characteristics in a Centrifugal Pump with Vaned Diffuser,” *Shock Vib.*, vol. 2018, pp. 1–11, Dec. 2018, doi: 10.1155/2018/9486536.

[26] H. Sun, S. Yuan, and Y. Luo, “Cyclic Spectral Analysis of Vibration Signals for Centrifugal Pump Fault Characterization,” *IEEE Sens. J.*, vol. 18, no. 7, pp. 2925–2933, Apr. 2018, doi: 10.1109/JSEN.2018.2804908.

[27] R. Kiliç, “Determination of Imbalance Problem in Electric Motor and Centrifugal Pump by Vibration Analysis,” *Acta Phys. Pol. A*, vol. 130, no. 1, pp. 487–491, Jul. 2016, doi: 10.12693/APhysPolA.130.487.

[28] C. Scheffer, “Pump Condition Monitoring Through Vibration Analysis,” *PUMPS Maint. Des. Reliab. Conf. 2008 - IDC Technol.*, pp. 1–20, 2008.

[29] M. Abd-Elaal, M. Lotfy, T. Nabil, and A. Eldomiaty, “Condition-based monitoring of a small centrifugal pump by vibration analysis,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 610, p. 012049, Oct. 2019, doi: 10.1088/1757-899X/610/1/012049.

[30] A. M. Abdulaziz and A. Kotb, “Detection of pump cavitation by vibration signature,” *Aust. J. Mech. Eng.*, vol. 15, no. 2, pp. 103–110, May 2017, doi: 10.1080/14484846.2015.1093261.

[31] A. Jami and P. S. Heyns, “Impeller fault detection under variable flow conditions based on three feature extraction methods and artificial neural networks,” *J. Mech. Sci. Technol.*, vol. 32, no. 9, pp. 4079–4087, Sep. 2018, doi: 10.1007/s12206-018-0807-3.

[32] A. K. Panda, J. S. Rapur, and R. Tiwari, “Prediction of flow blockages and impending cavitation in centrifugal pumps using Support Vector Machine (SVM) algorithms based on vibration

measurements," *Measurement*, vol. 130, pp. 44–56, Dec. 2018, doi: 10.1016/j.measurement.2018.07.092.

[33] V. Muralidharan and V. Sugumaran, "SELECTION OF DISCRETE WAVELETS FOR FAULT DIAGNOSIS OF MONOBLOCK CENTRIFUGAL PUMP USING THE J48 ALGORITHM," *Appl. Artif. Intell.*, vol. 27, no. 1, pp. 1–19, Jan. 2013, doi: 10.1080/08839514.2012.721694.

[34] V. Muralidharan and V. Sugumaran, "A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis," *Appl. Soft Comput.*, vol. 12, no. 8, pp. 2023–2029, Aug. 2012, doi: 10.1016/j.asoc.2012.03.021.

[35] A. Daraz, S. Alabied, F. Gu, and A. D. Ball, "Modulation Signal Bispectrum Analysis of Acoustic Signals for the Impeller Wear Detection of Centrifugal Pumps," in *2019 25th International Conference on Automation and Computing (ICAC)*, 2019, pp. 1–6.

[36] N. Zhang, M. Yang, B. Gao, Z. Li, and D. Ni, "Experimental and numerical analysis of unsteady pressure pulsation in a centrifugal pump with slope volute," *J. Mech. Sci. Technol.*, vol. 29, no. 10, pp. 4231–4238, Oct. 2015, doi: 10.1007/s12206-015-0919-y.

[37] Ono Sokki Co., Ltd., "VC 2200 Vibration Comparator Specification," 2018. www.onosokki.co.jp.

[38] SPM Instruments, "Intellinova HD-ENV," 2019. https://www.spminstrument.com/Products/.

[39] M. T. Shervani-Tabar, M. M. Ettefagh, S. Lotfan, and H. Safarzadeh, "Cavitation intensity monitoring in an axial flow pump based on vibration signals using multi-class support vector machine," *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 232, no. 17, pp. 3013–3026, Sep. 2018, doi: 10.1177/0954406217729416.

[40] B. Bohn, R. Khoie, R. B. Gopaluni, J. A. Olson, and B. Stoeber, "Development and Characterization of a Non-Intrusive Sensor to Measure Wear in Centrifugal Pumps," *IEEE Sens. J.*, vol. 19, no. 18, pp. 7906–7914, Sep. 2019, doi: 10.1109/JSEN.2019.2919283.

[41] T. Ahonen, J. T. Kortelainen, J. K. Tamminen, and J. Ahola, "Centrifugal pump operation monitoring with motor phase current measurement," *Int. J. Electr. Power Energy Syst.*, vol. 42, no. 1, pp. 188–195, Nov. 2012, doi: 10.1016/j.ijepes.2012.04.013.

[42] S. Yuan, J. Yang, J. Yuan, Y. Luo, and J. Pei, "Experimental investigation on the flow-induced noise under variable conditions for centrifugal pumps," *Chin. J. Mech. Eng.*, vol. 25, no. 3, pp. 456–462, May 2012, doi: 10.3901/CJME.2012.03.456.

[43] A. Suhane, "Experimental Study on Centrifugal Pump to Determine the Effect of Radial Clearance on Pressure Pulsations, Vibrations and Noise," *Int. J. Eng. Res. Appl.*, vol. 2, no. 4, pp. 1823–1829, Aug. 2012.

[44] Honeywell, "Heavy Duty Pressure Transducers Data Sheet - 32328895." Honeywell International Inc., 2018.

[45] Measurement Specialties Inc., "M5200 Industrial Pressure Transducer Data Sheet." TE Connectivity, 2020.

[46] S. Li, N. Chu, P. Yan, D. Wu, and J. Antoni, "Cyclostationary approach to detect flow-induced effects on vibration signals from centrifugal pumps," *Mech. Syst. Signal Process.*, vol. 114, pp. 275–289, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.027.

[47] E. H. Higham and S. Perovic, “Predictive maintenance of pumps based on signal analysis of pressure and differential pressure (flow) measurements,” *Trans. Inst. Meas. Control*, vol. 23, no. 4, pp. 226–248, 2001.

[48] T. Xie, S. M. Ghiaasiaan, and S. Karrila, “Flow Regime Identification in Gas/Liquid/Pulp Fiber Slurry Flows Based on Pressure Fluctuations Using Artificial Neural Networks,” *Ind. Eng. Chem. Res.*, vol. 42, no. 26, pp. 7017–7024, Dec. 2003, doi: 10.1021/ie0304199.

[49] Sulzer Pumps, “Vertical Pump Troubleshooting Guide.” Sulzer Pumps, 2012.

[50] P. M. Bardet, P. F. Peterson, and ö. Savaş, “Annular swirling liquid layer with a hollow core,” *J. Fluid Mech.*, vol. 841, pp. 784–824, Apr. 2018, doi: 10.1017/jfm.2018.76.

[51] T. Schäfer, M. Neumann, A. Bieberle, and U. Hampel, “Experimental investigations on a common centrifugal pump operating under gas entrainment conditions,” *Nucl. Eng. Des.*, vol. 316, pp. 1–8, May 2017, doi: 10.1016/j.nucengdes.2017.02.035.

[52] M. Stan, I. Pana, M. Minescu, A. Ichim, and C. Teodoriu, “Centrifugal Pump Monitoring and Determination of Pump Characteristic Curves Using Experimental and Analytical Solutions,” *Processes*, vol. 6, no. 2, p. 18, Feb. 2018, doi: 10.3390/pr6020018.

[53] Q. Si, W. He, G. Bois, Q. Cui, S. Yuan, and K. Zhang, “Experimental and Numerical Studies on Flow Characteristics of Centrifugal Pump under Air-water Inflow,” *Int. J. Fluid Mach. Syst.*, vol. 12, no. 1, pp. 31–38, Mar. 2019, doi: 10.5293/IJFMS.2019.12.1.031.

[54] Y. Xu, S. Cao, T. Sano, T. Wakai, and M. Reclari, “Experimental Investigation on Transient Pressure Characteristics in a Helico-Axial Multiphase Pump,” *Energies*, vol. 12, no. 3, p. 461, Jan. 2019, doi: 10.3390/en12030461.

[55] T. Xie, S. M. Ghiaasiaan, and S. Karrila, “Artificial neural network approach for flow regime classification in gas–liquid–fiber flows based on frequency domain analysis of pressure signals,” *Chem. Eng. Sci.*, vol. 59, no. 11, pp. 2241–2251, Jun. 2004, doi: 10.1016/j.ces.2004.02.017.

[56] Y. A. Khalid and S. M. Sapuan, “Wear analysis of centrifugal slurry pump impellers,” *Ind. Lubr. Tribol.*, vol. 59, no. 1, pp. 18–28, Feb. 2007, doi: 10.1108/00368790710723106.

[57] Y. Wang, M. J. Zuo, and X. Fan, “Design of an experimental system for wear assessment of slurry pumps,” *Proc. Can. Eng. Educ. Assoc. CEEA*, 2005.

[58] M. E. Matlakala, D. V. V. Kallon, K. E. Mogapi, I. M. Mabelane, and D. M. Makgopa, “Influence of Impeller Diameter on the Performance of Centrifugal pumps.,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 655, p. 012009, Nov. 2019, doi: 10.1088/1757-899X/655/1/012009.

[59] L. Cao, Z. Wang, Y. Xiao, and Y. Luo, “Numerical Investigation of Pressure Fluctuation Characteristics in a Centrifugal Pump with Variable Axial Clearance,” *Int. J. Rotating Mach.*, vol. 2016, pp. 1–11, 2016, doi: 10.1155/2016/9306314.

[60] Z. Yao, F. Wang, L. Qu, R. Xiao, C. He, and M. Wang, “Experimental Investigation of Time-Frequency Characteristics of Pressure Fluctuations in a Double-Suction Centrifugal Pump,” *J. Fluids Eng.*, vol. 133, no. 10, p. 101303, Oct. 2011, doi: 10.1115/1.4004959.

[61] M. Tan, J. Feng, H. Liu, J. Ding, and Z. Zhu, "Experimental Investigation of Unsteady Characteristics in a Single-Channel Pump," *J. Test. Eval.*, vol. 47, no. 1, p. 20170330, Jan. 2019, doi: 10.1520/JTE20170330.

[62] Ampco Pumps Company, "AC-114 Series Data Sheet." Ampco Pumps Company, 2007.

[63] T. Shih, W. W. Liou, A. Shabbir, Z. Yang, and J. Zhu, "A New k-ε Eddy Viscosity Model for High Reynolds Number Turbulent Flows," *Comput. Fluids*, vol. 24, no. 3, pp. 227–238, 1995.

[64] S. Huang, X. Su, J. Guo, and L. Yue, "Unsteady numerical simulation for gas–liquid two-phase flow in self-priming process of centrifugal pump," *Energy Convers. Manag.*, vol. 85, pp. 694–700, Sep. 2014, doi: 10.1016/j.enconman.2014.03.023.

[65] K. Minemura and T. Uchiyama, "Three-Dimensional Calculation of Air-Water Two-Phase Flow in Centrifugal Pump Impeller Based on a Bubbly Flow Model," *J. Fluids Eng.*, vol. 115, no. 4, pp. 766–771, Dec. 1993, doi: 10.1115/1.2910210.

[66] E. Olsson and G. Kreiss, "A conservative level set method for two phase flow," *J. Comput. Phys.*, vol. 210, no. 1, pp. 225–246, Nov. 2005, doi: 10.1016/j.jcp.2005.04.007.

[67] Rosemount Inc., "Rosemount 8700 Series Magnetic Flowmeter Sensors - Reference Manual 00809-0100-4727, Rev EA." 2010.

[68] Measurement Computing Corporation, "USB-1208HS Series Data Sheet, Rev 6." Measurement Computing Corporation, 2018.

[69] D. Zwillinger and S. Kokoska, *CRC Standard Probability and Statistics Tables and Formulae*. CRC Press, 1999.

[70] SciPy 1.0 Contributors *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nat. Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020, doi: 10.1038/s41592-019-0686-2.

[71] A. Lenail, "NN-SVG Visualizer," 2019. http://alexlenail.me/NN-SVG/index.html.

[72] C. Beleites, R. Salzer, and V. Sergo, "Validation of Soft Classification Models using Partial Class Memberships: An Extended Concept of Sensitivity & Co. applied to the Grading of Astrocytoma Tissues," *Chemom. Intell. Lab. Syst.*, vol. 122, pp. 12–22, Mar. 2013, doi: 10.1016/j.chemolab.2012.12.003.

[73] S. Nakatani, K. Yoshida, and Y. Watanabe, "Optical and Acoustic Observations of Bubble Adhered to Piezoelectric Transducer under Ultrasound Field: Pressure Signal Caused by Bubble Behavior," *Jpn. J. Appl. Phys.*, vol. 47, no. 5, pp. 3904–3907, May 2008, doi: 10.1143/JJAP.47.3904.

[74] S. Chu, R. Dong, and J. Katz, "Relationship between unsteady flow, pressure fluctuations, and noise in a centrifugal pump—Part A: Use of PDV data to compute the pressure field," 1995.

APPENDIX A – SUPPLEMENTARY CALCULATIONS

A1 Turbulence Parameters

The mean strain rate tensor

$$\boldsymbol{S} = \frac{1}{2}\left[\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T\right] . \tag{A1}$$

Following, the strain rate tensor

$$\mathrm{S} = \sqrt{2\boldsymbol{S} : \boldsymbol{S}} . \tag{A2}$$

The turbulent kinetic energy production term

$$P_k = \mu_{\mathrm{Tl}}\left[\nabla\boldsymbol{u} : (\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T) - \frac{2}{3}(\nabla\cdot\boldsymbol{u})^2\right] - \frac{2}{3}\rho k_{\mathrm{l}}\nabla\cdot\boldsymbol{u} . \tag{A3}$$

The model constant

$$C_\mu = \frac{1}{A_o + A_S U^{(*)}\frac{k_{\mathrm{l}}}{\epsilon}} , \tag{A4}$$

where the associated constants

$$A_o = 4 , \tag{A5}$$

$$A_S = \sqrt{6}\ \cos\left(\frac{1}{3}\cos^{-1}\left(\sqrt{6}W\right)\right) , \tag{A6}$$

$$W = \frac{2\sqrt{2}\boldsymbol{S}:(\boldsymbol{S}\cdot\boldsymbol{S})}{|\boldsymbol{S}|^3} , \tag{A7}$$

and

$$U^{(*)} = \sqrt{\boldsymbol{S} : \boldsymbol{S} + \boldsymbol{\Omega} : \boldsymbol{\Omega}} . \tag{A8}$$

In (A8), the mean rotation rate tensor

$$\boldsymbol{\Omega} = \frac{1}{2}\left[\nabla\boldsymbol{u} - (\nabla\boldsymbol{u})^T\right] . \tag{A9}$$

A2 Pearson Correlation Coefficient

For two signals, x and y, each with N samples, the Pearson correlation coefficient

$$r_{\mathrm{P}} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}\,. \tag{A10}$$

APPENDIX B – SUPPLEMENTARY FIGURES

B1 Virtual Probe Placement Dependency

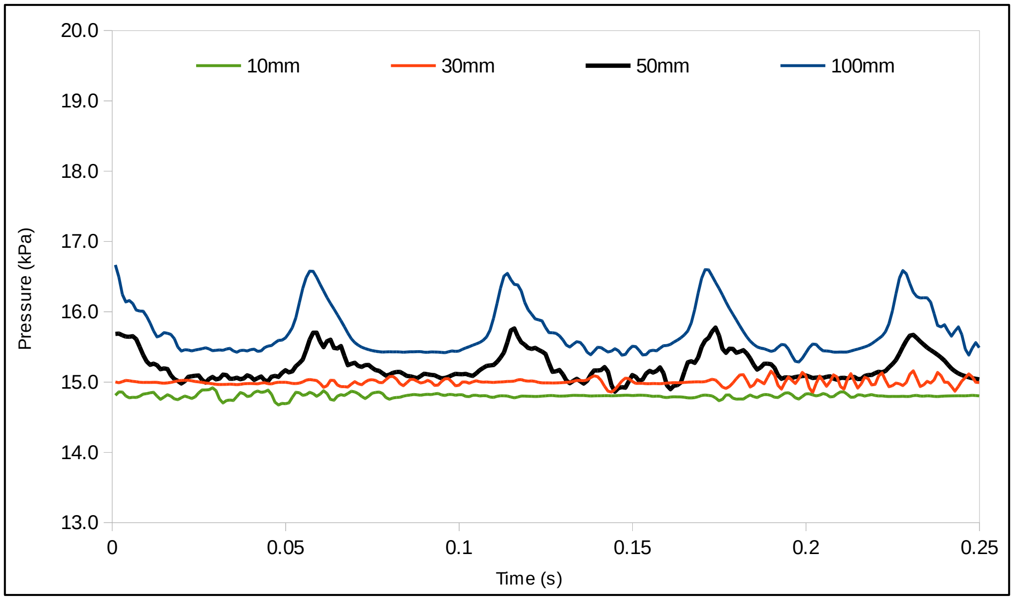


Figure B1: Simulated pressure fluctuations as a function of virtual probe placement (in mm from the model outlet boundary) at $\boldsymbol{\omega = 8.75Hz}$. The black line (50 mm) represents the selected position.

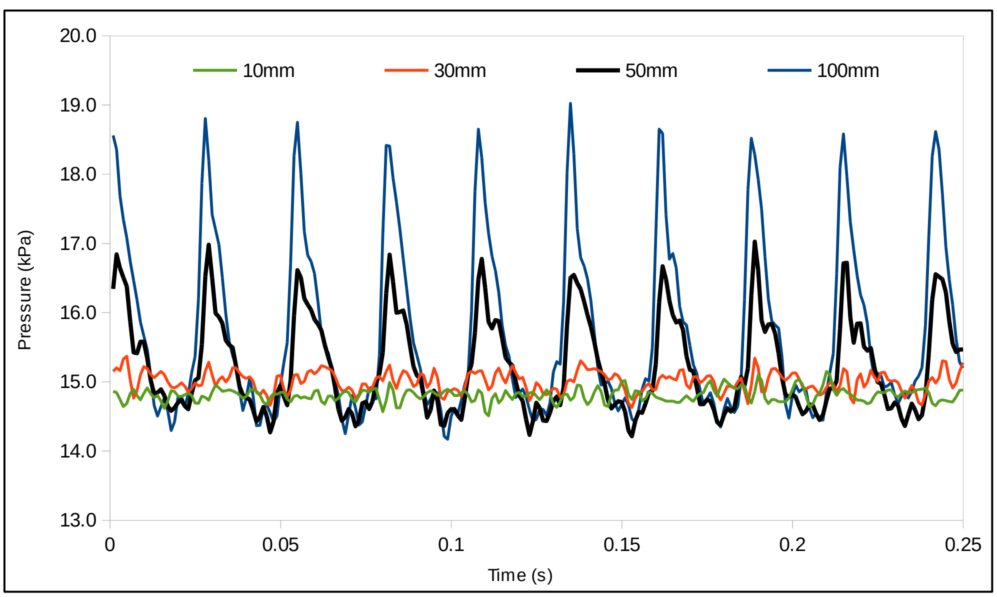


Figure B2: Simulated pressure fluctuations as a function of virtual probe placement (in mm from the model outlet boundary) at $\boldsymbol{\omega = 18.75Hz}$. The black line (50 mm) represents the selected position.

B2 Gas Entrainment Feature Optimization Plots (Experimental)

The following figures show the respective correlation of each feature to ω, then the resulting feature trend of its optimized form.

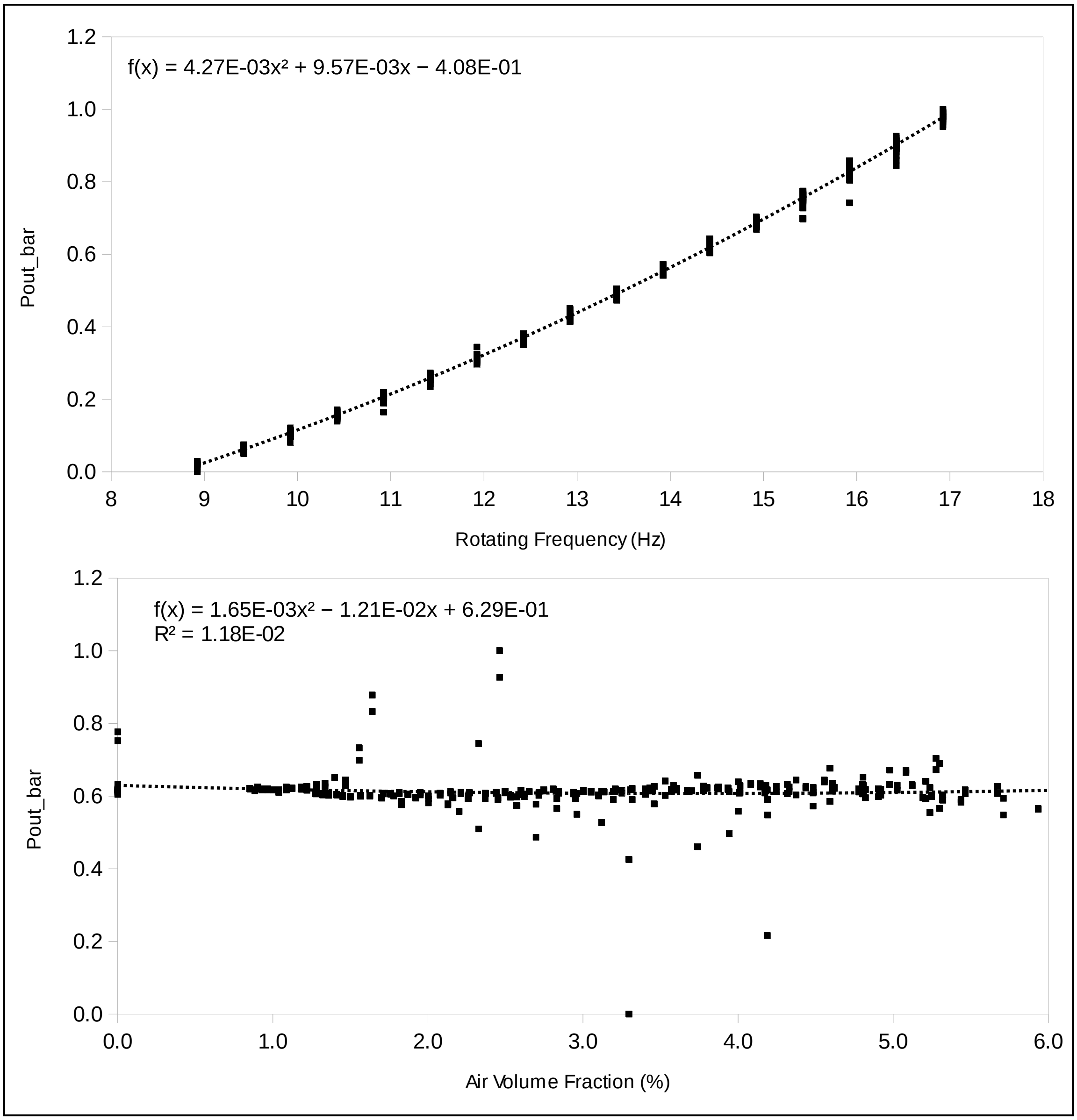


Figure B3: Optimization of $\overline{p_{\mathbf{out}}}$ for gas entrainment.

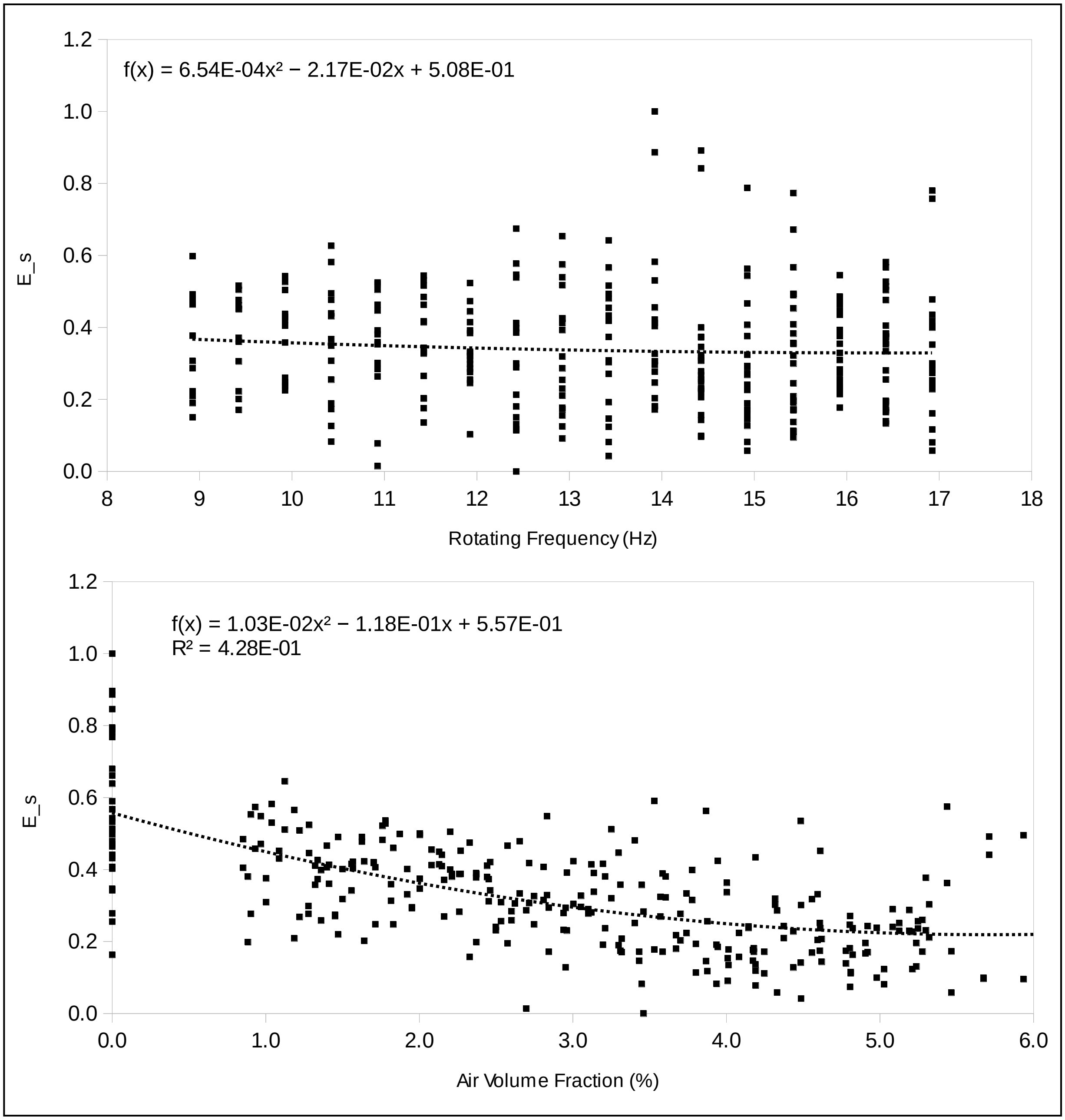


Figure B4: Optimization of E_s for gas entrainment.

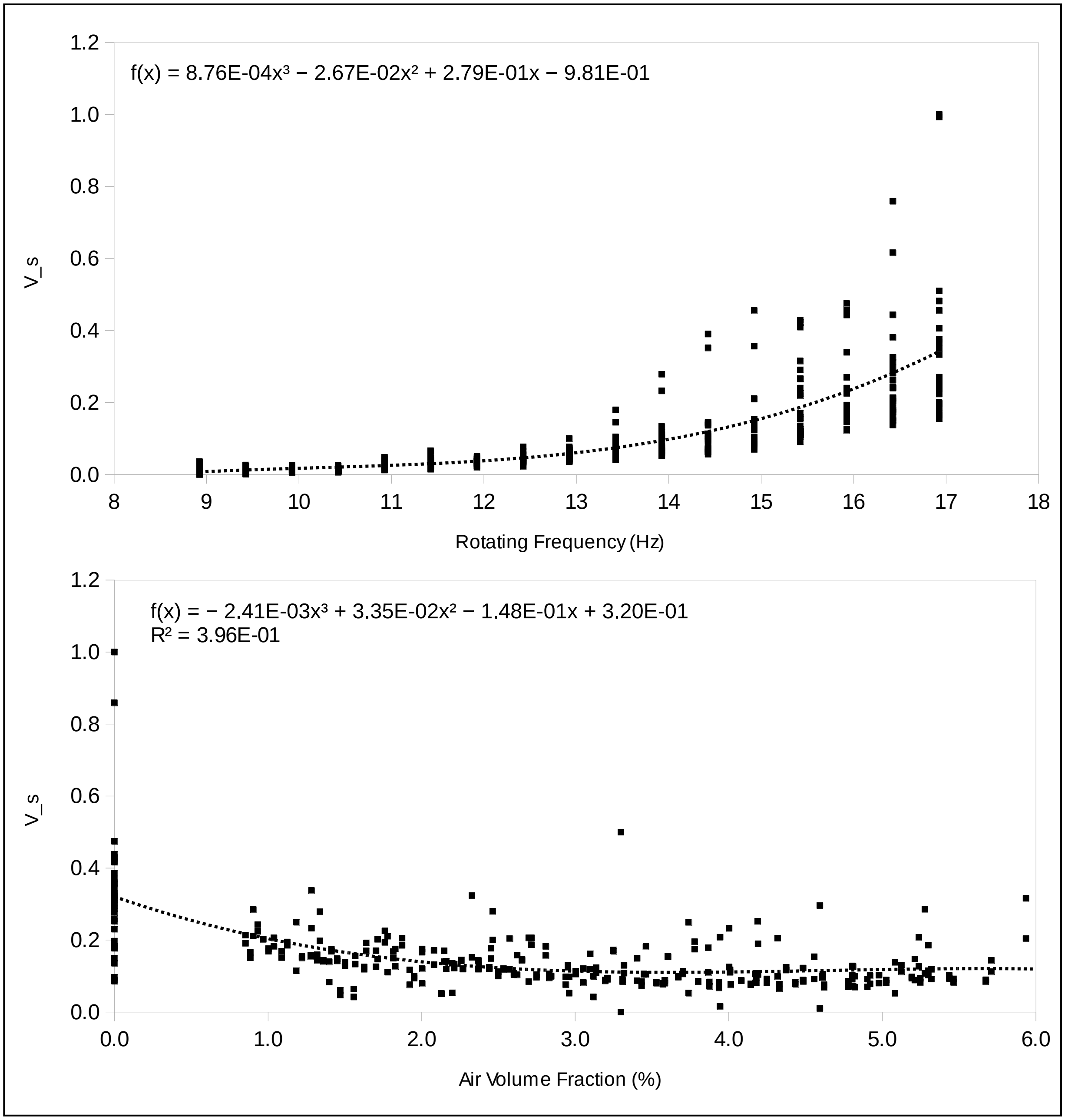


Figure B5: Optimization of V_s for gas entrainment.

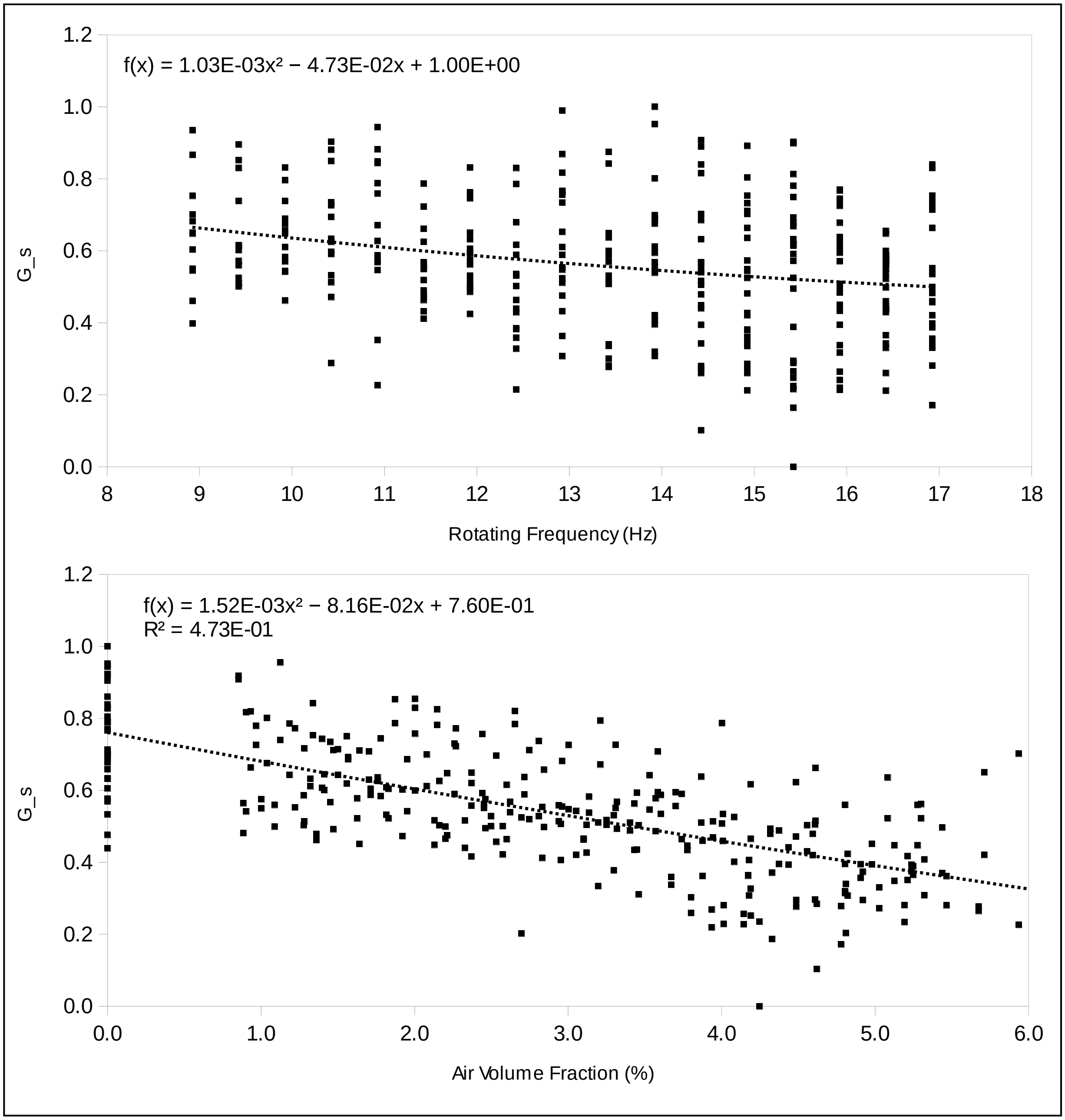


Figure B6: Optimization of G_s for gas entrainment.

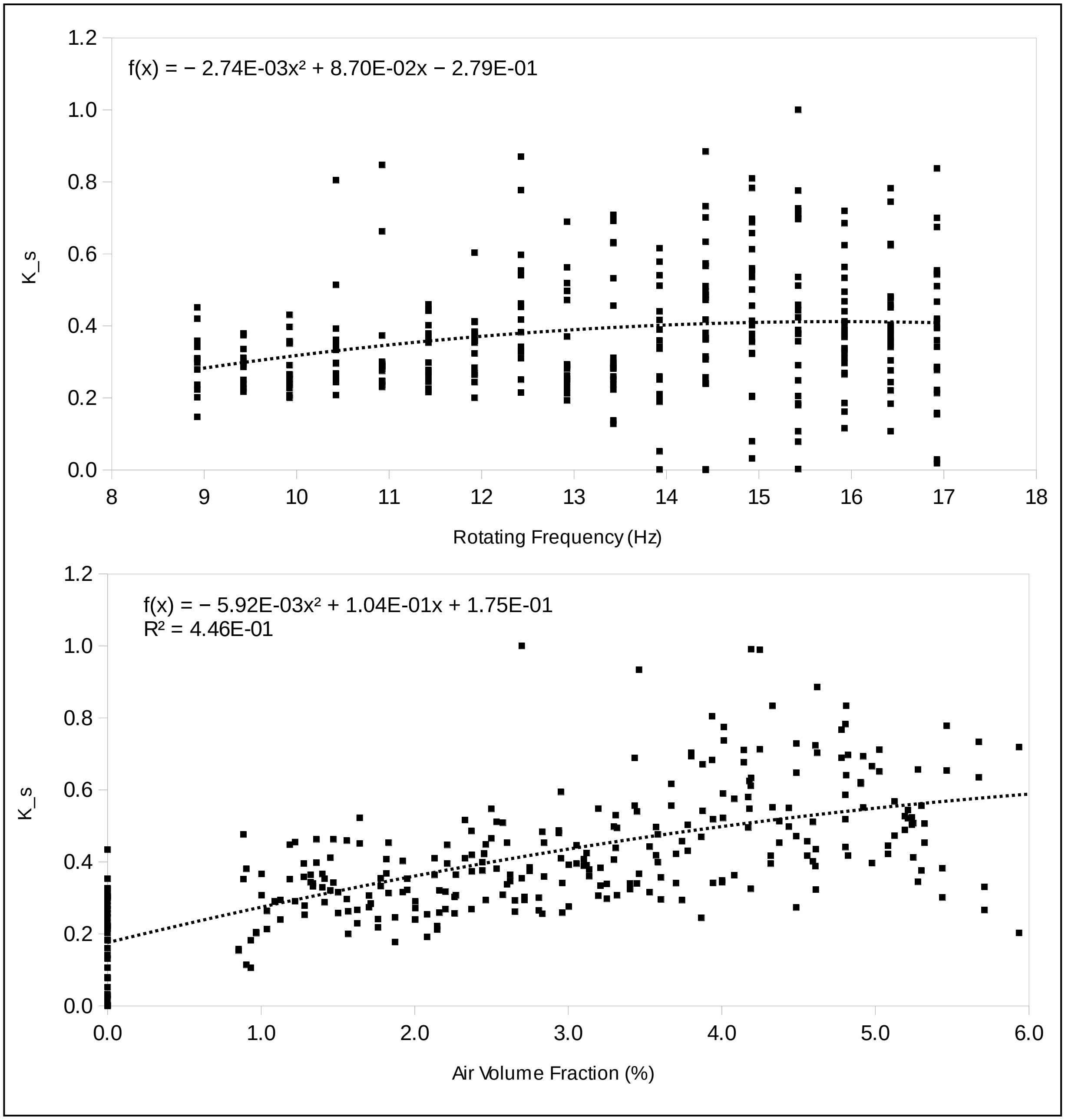


Figure B7: Optimization of K_s for gas entrainment.

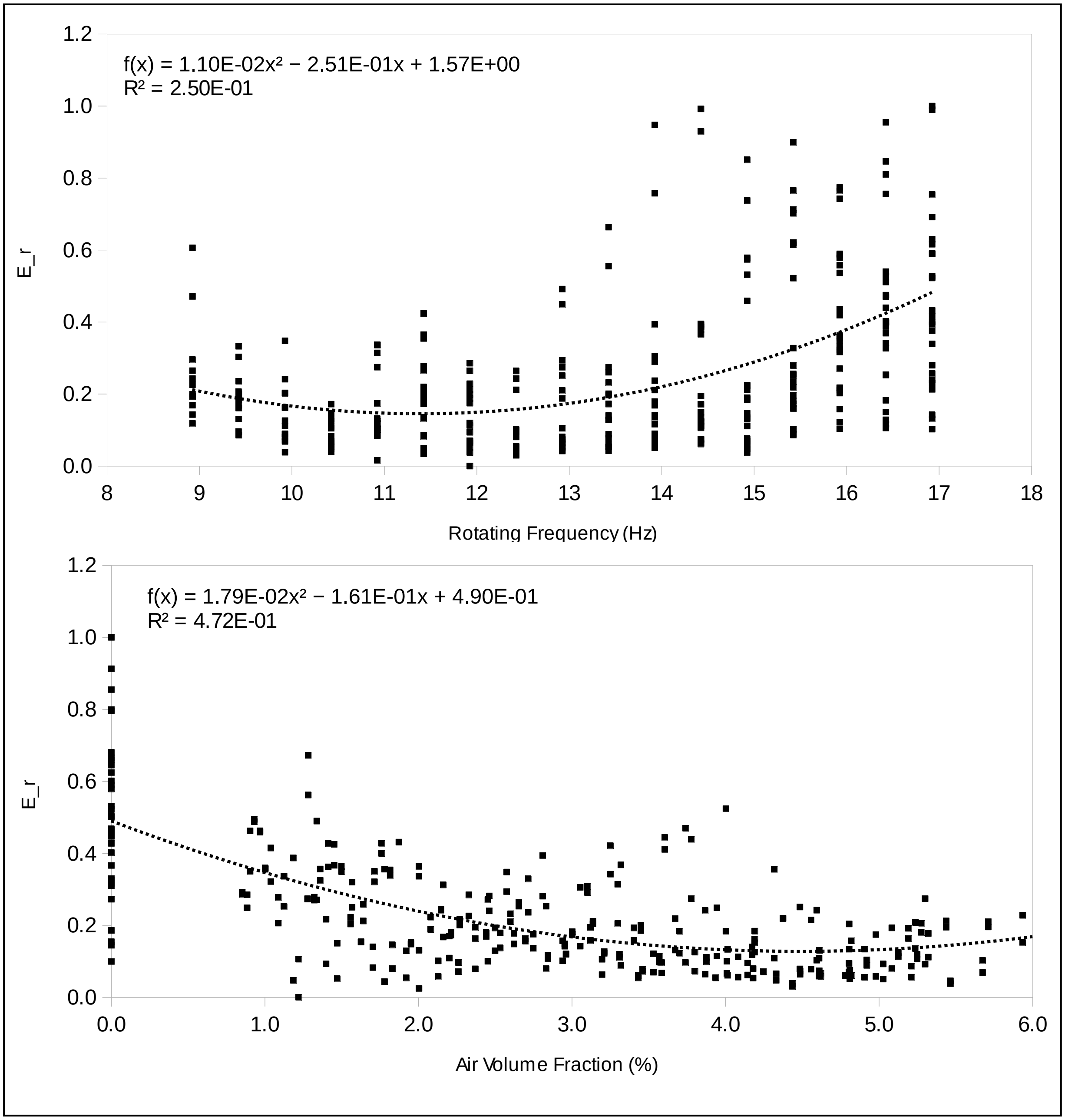


Figure B8: Optimization of $\boldsymbol{E}_{\mathbf{R}}$ for gas entrainment.

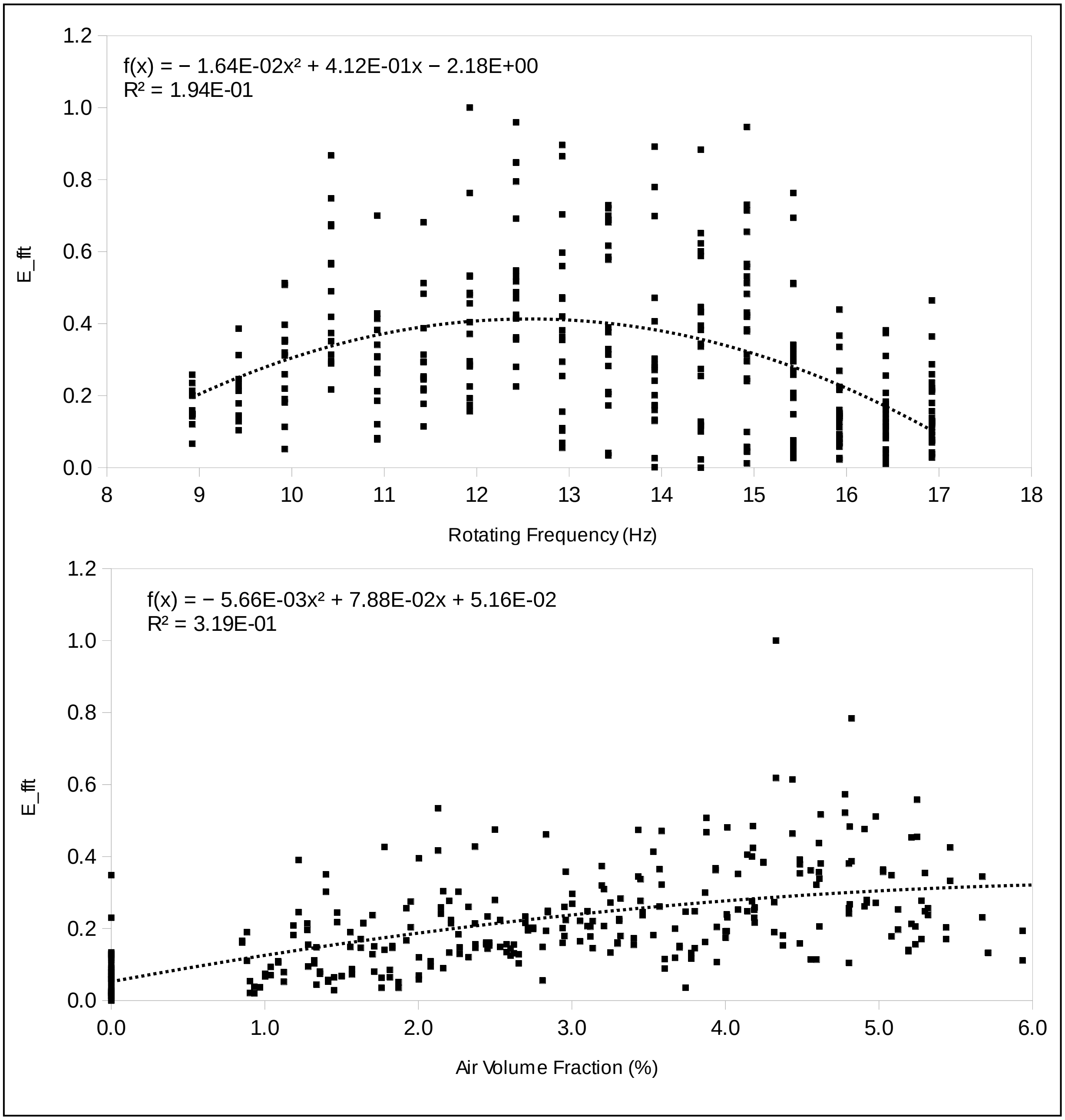


Figure B9: Optimization of $E_{\mathbf{FFT}}$ for gas entrainment.

B3 Impeller Wear Feature Optimization Plots (Simulated)

The following figures show the respective correlation of each feature to ω, then the resulting feature trend of its optimized form.

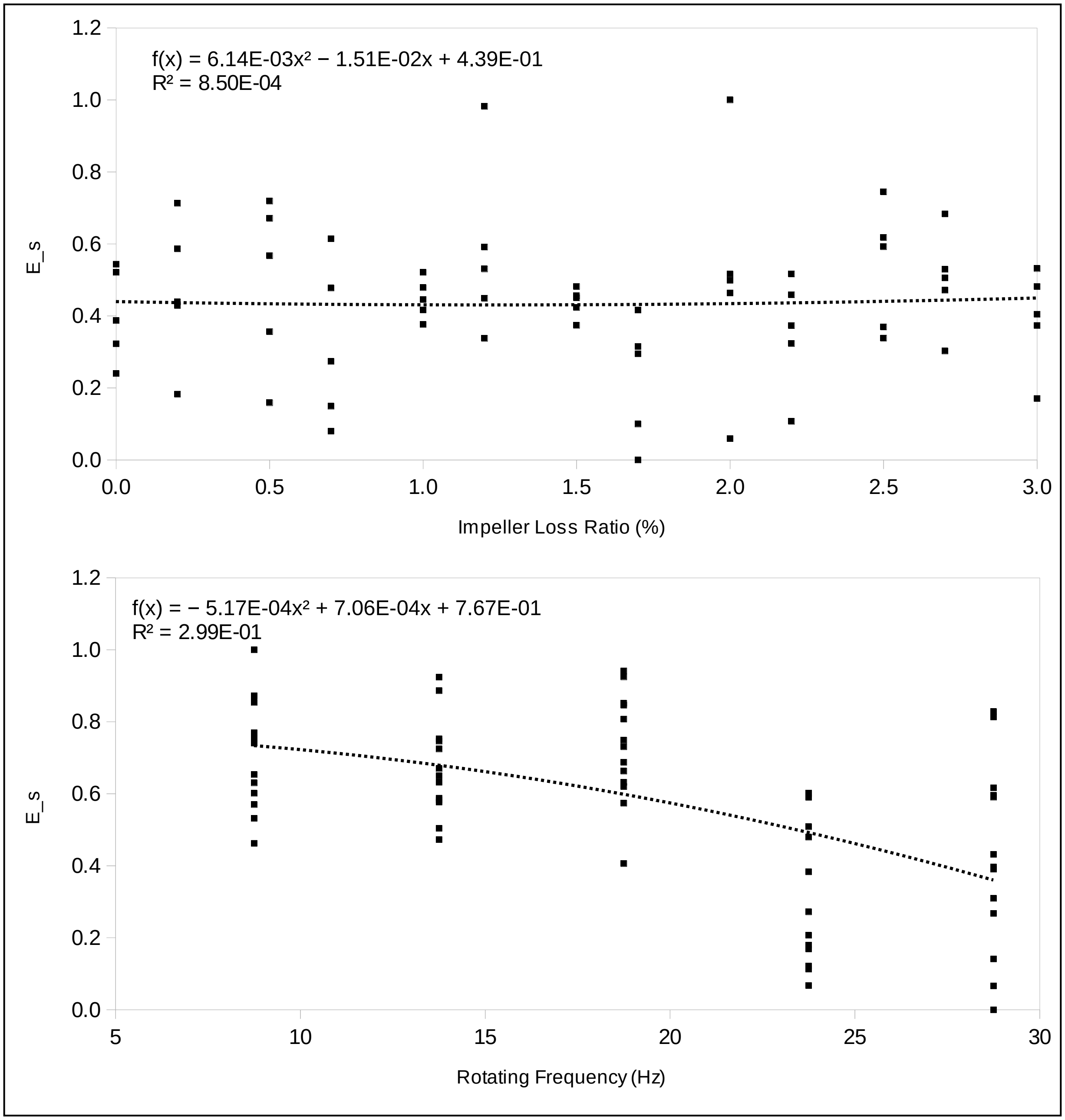


Figure B10: Optimization of $\boldsymbol{E_s}$ for impeller wear.

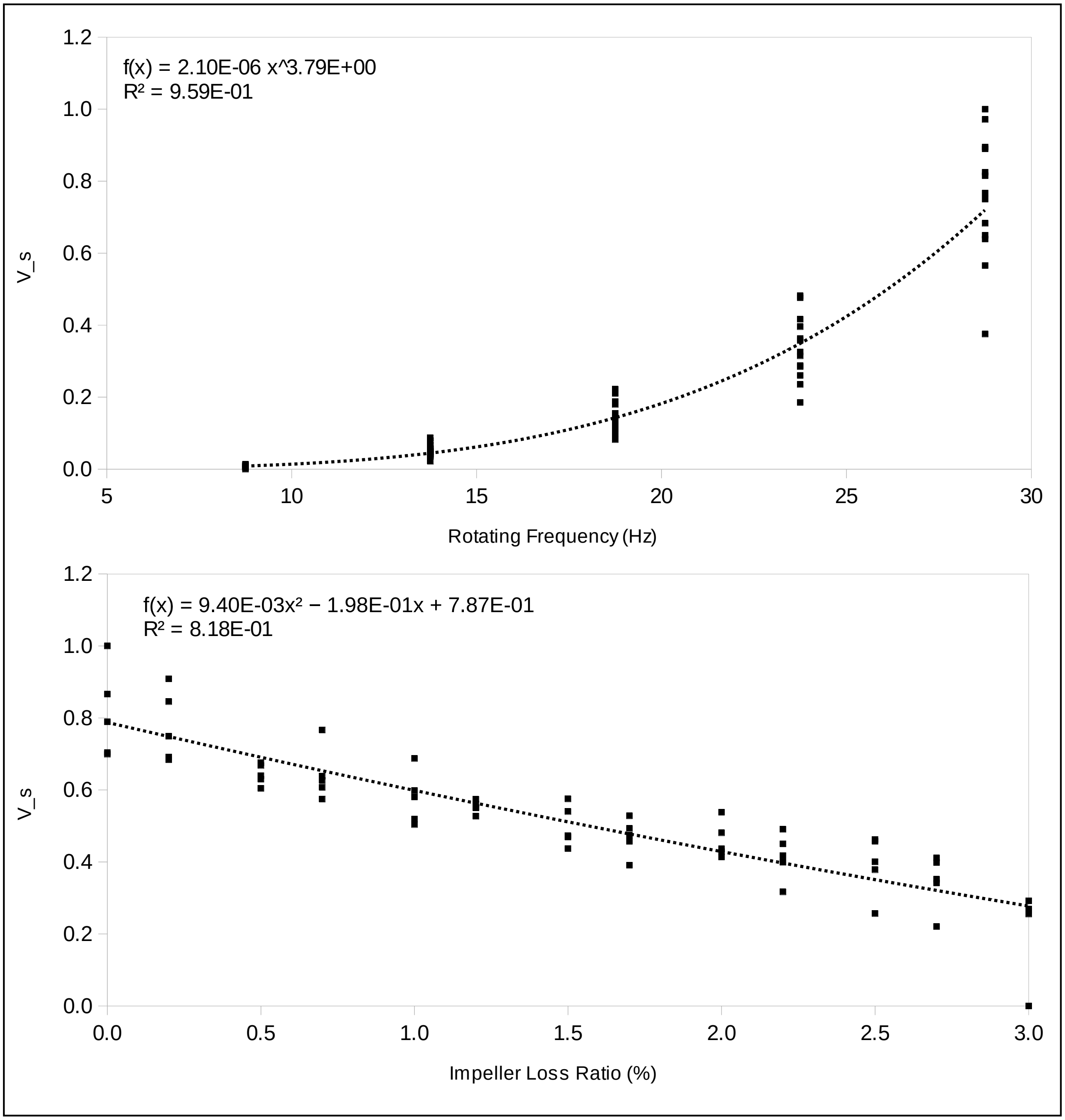


Figure B11: Optimization of V_s for impeller wear.

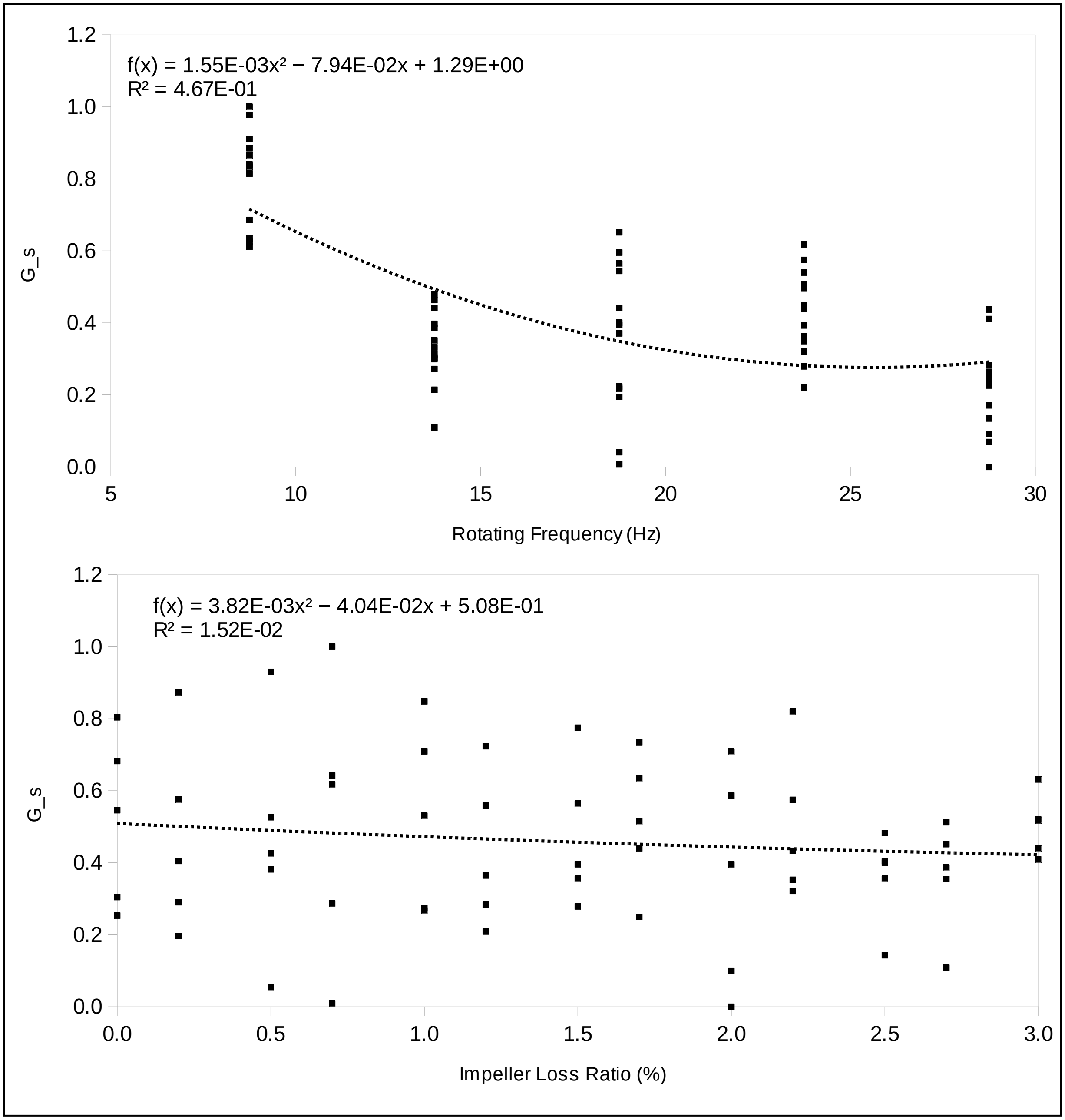


Figure B12: Optimization of G_s for impeller wear.

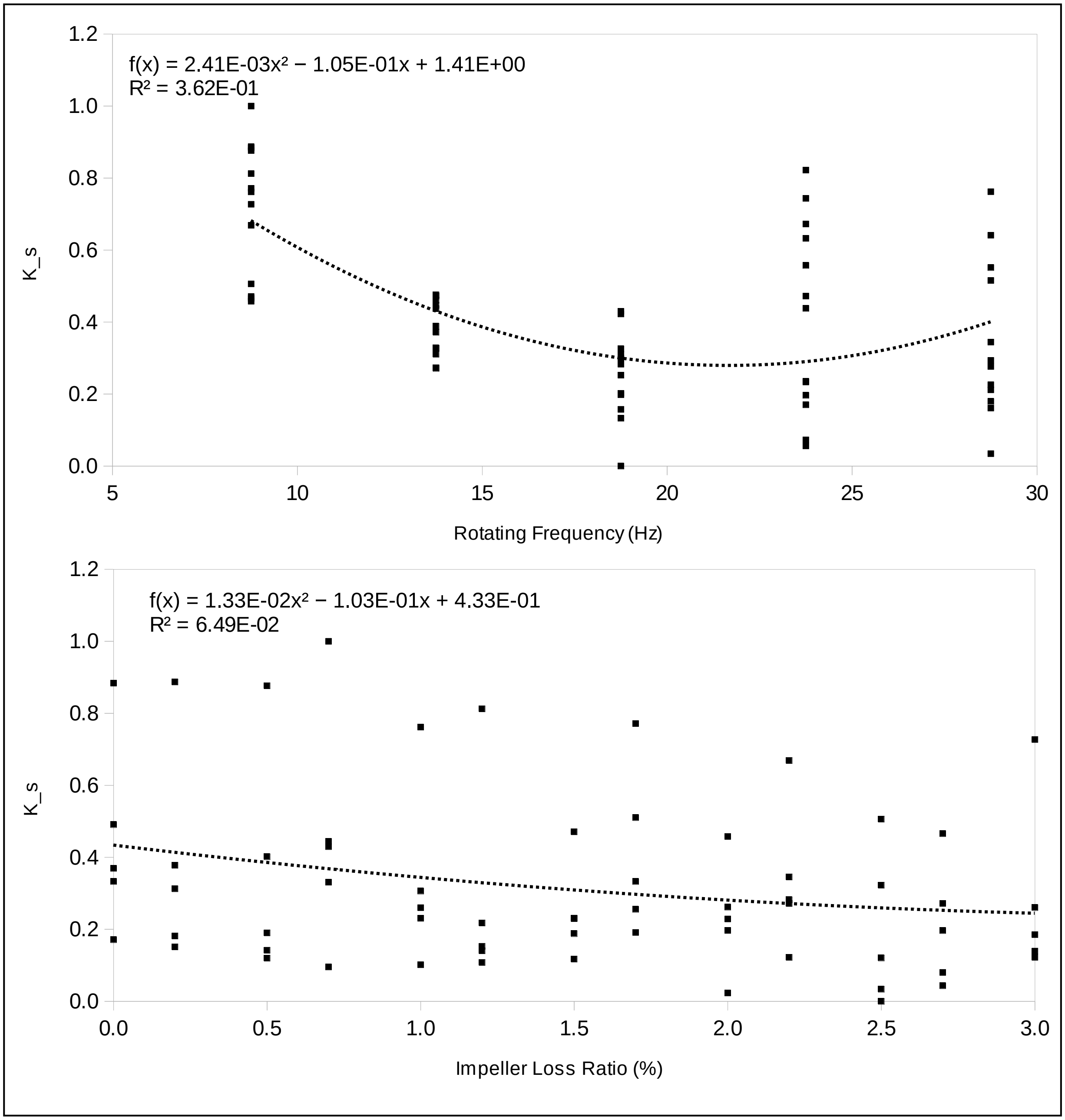


Figure B13: Optimization of K_s for impeller wear.

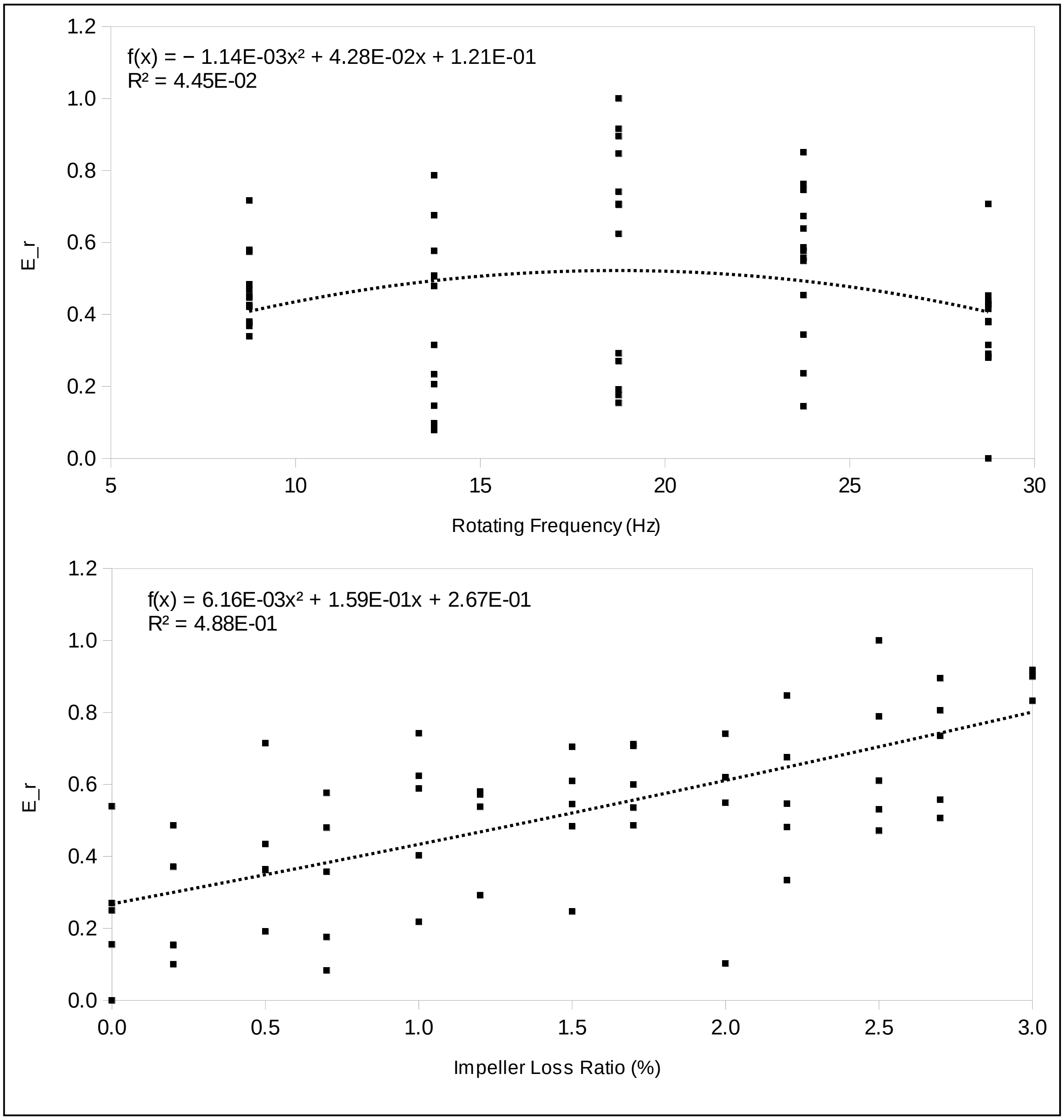


Figure B14: Optimization of E_R for impeller wear.

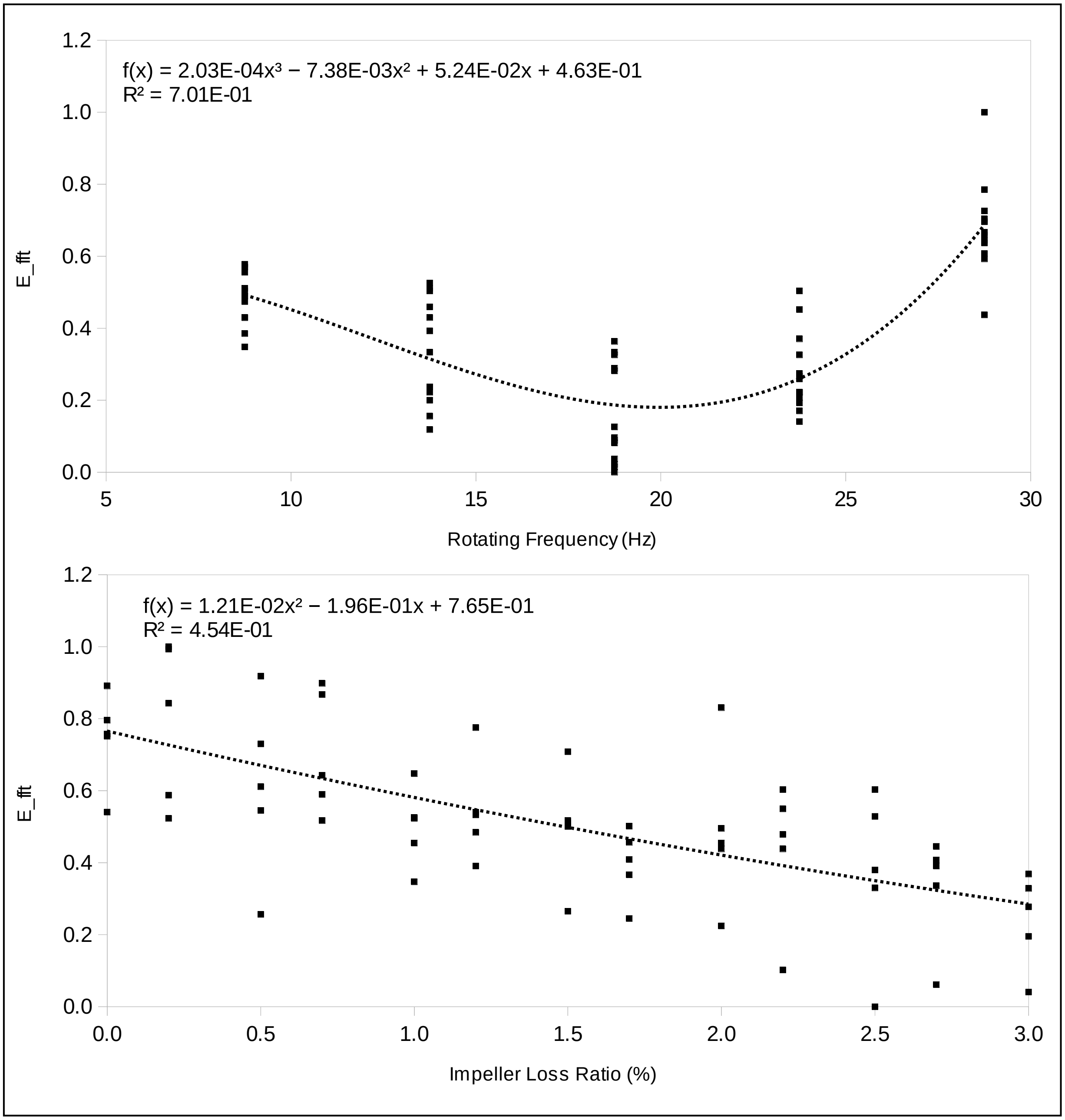


Figure B15: Optimization of E_{FFT} for impeller wear.