DATA-DRIVEN MODELLING OF MULTIPHASE FLOW SYSTEMS

Keivan Mokhtarpour

A thesis
in
The Department
of
Mechanical, Industrial and Aerospace Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Applied Science (Mechanical Engineering)
Concordia University
Montréal, Québec, Canada

September 2020

Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: **Keivan Mokhtarpour**
Entitled: **Data-Driven Modelling of Multiphase Flow Systems**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Mechanical Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

_____ Chair
Lyes Kadem

_____ Examiner
Lyes Kadem

_____ Examiner
Samuel Li

_____ Thesis Supervisor
Ali Dolatabadi

Approved _____
Mamoun Medraj, Graduate Program Director

Date _____
25/08/2020 Mourad Debbabi, Dean of Faculty

Abstract

Data-Driven Modelling of Multiphase Flow Systems

Keivan Mokhtarpour

Dynamical systems specifically in the field of fluid mechanics are composed of underlying complicated governing phenomena originated from nonlinearities and instabilities. Encountered with the challenge of analyzing vast amount of data, the concept of reduced order modelling (ROM) was emerged to map the high resolution spatio-temporal data onto a low-dimensional space using the most prominent embedded features. This dissertation considers two ROM techniques of proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD) applied to liquid injection systems. These approaches have been widely used to tackle the challenges of analyzing spatio-temporal coherence of dynamical systems. Despite the numerous works implementing POD and DMD, there has been a lack of physical meaning for the modes generated by them. An interpretation of POD and DMD modes is provided in this thesis by the recognition of dominating features. The main focus will be primitively on benchmark problems to validate the efficacy of the methods and consequently to the liquid jets exposed to air crossflows in a hierarchical scheme. A grasp of the prominent spatial structures and their corresponding leading dynamic frequencies will be provided through the analysis of POD and DMD frequency spectra. Effects of several different factors such as the gaseous Weber number, liquid-gas momentum flux ratio and the injector aspect ratio are investigated in this study. Finally, the power of ROM techniques to create features for machine-learnt classifiers that are sufficient for categorization of sundry types of flow regimes is investigated in a supervised manner. These classifiers are opted from a range of classical machine learning algorithms like support vector machines (SVM) and random forest (RF) that have been extensively employed for classification tasks in the recent years. The best combination of reduced order models with the machine learning algorithms are presented.

Acknowledgments

I would like to express my veridical gratitude to my supervisor, Prof. Ali Dolatabadi, who continuously guided and patronized me even at times when the road was bumpy. Without his persistent help, the completion of this project would not have been achieved.

The financial support provided by the National Sciences and Engineering Research Council (NSERC) is truly appreciated. Without their support, this work could not have reached its goal.

I also wish to acknowledge the great love of my family, my mother, Nura; my father, Hossein; and my brother, Kiavash. They kept me going from the very beginning with their positiveness.

At the end, I am very thankful to my dear friend and colleague Dr. Mehdi Jadidi for all the technical support he provided me with whenever I was encountered with an issue.

Contents

List of Figures

List of Tables

List of Symbols

$\mathbb{A}$	Matrix of discrete-time linear dynamics
$\tilde{A}$	Reduced dynamics on POD subspace
A	Matrix of continuous-time linear dynamics
Ar	Aspect ratio
D	Cylinder/Nozzle diameter
E	Energy content definition
f	Frequency
f_s	Sampling frequency
l	Characteristic length
m	Number of data snapshots
N	Nonlinear partial differential equation
n	Dimension of the spatial domain
q	Momentum flux ratio
Re	Reynolds number
r	Rank of truncated SVD
S	Mode's stability definition
St	Strouhal number
U	Left singular vectors (POD modes) of X
u	Velocity
u_∞	Free-stream fluid velocity
V	Right singular vectors of X (matrix of temporal vectors)
W	Eigenvectors of $\tilde{A}$
We	Gaseous Weber number
X	Data matrix
x_k	Snapshot of data at time t_k

Greek Symbols

Λ	Diagonal matrix of DMD eigenvalues (discrete-time)
λ	DMD eigenvalue
Φ	Matrix of DMD modes
ϕ	DMD mode
Σ	Matrix of singular values / Summation
σ	Surface tension
σ_k	(k)th singular value
ν	Kinematic viscosity of fluid
ρ	Density of fluid

Acronyms

CFD	Computational fluid dynamics
DMD	Dynamic mode decomposition
FFT	Fast Fourier transform
GUI	Graphical user interface
LJIC	Liquid jet in crossflow
PCA	Principal component analysis
PIV	Particle image velocimetry
POD	Proper orthogonal decomposition
PSD	Power spectral density
RF	Random forest
ROM	Reduced order modelling
SVD	Singular value decomposition
SVM	Support vector machine

Preface

The content of this thesis has been organized in a manuscript style. The results of the third chapter has been previously accepted for the canadian society of mechanical engineering (CSME 2020) conference and was partially submitted to the transactions of the canadian society for mechanical engineering (TCSME) journal. The findings of the fourth chapter were also submitted to the journal of atomizations of sprays. Considering the fact that this thesis is organized in a manuscript based format, each chapter is dedicated to a specific discussion. Chapter one brings about a primary understanding of reduced order modelling techniques with a great focus on POD and DMD where relevant literature to the current work under study are addressed. The motivations for performing the current research and the existing challenges are then implied. Chapter two considers the underlying governing formulations of POD and DMD, based on which, two benchmark problems are analysed and the results get correlated to the existing base knowledge. A hierarchical investigation of liquid breakup systems is the next step in this chapter where a physical interpretation of such systems is debatable. In chapter three, DMD is applied to different cases of elliptical liquid jets exposed to air crossflows at low gaseous Weber numbers. A deep scrutiny of such systems is performed with a comparative discussion on the characteristics of circular and elliptical jets. Chapter four propounds the idea of using the POD/DMD modes as the reliable features to optimally classify different flow regimes of liquid jets in crossflow. Machine learning algorithms are employed to attain this goal. Chapter five finalizes the discussion on the previous chapters by deducing the pertaining results to the performed work and drawing the perspective ahead.

Chapter 1

Introduction

Data science is usually defined as an interdisciplinary field that uses methods, algorithms and processes to extract insights from structured or non-structured data. Machine learning (ML) as a well-known branch of data science is the study of computer algorithms that helps gaining knowledge from data. These algorithms improve automatically through experience and are widely applied on various fields such as engineering, finance, biology, neuroscience and etc. The interface of ML with the fluid mechanics has received a great attention over the last few years. Flow modelling, flow control and optimization tasks either in experiments or simulations have been always of the most challenging works of researchers and engineers where the amount of data to be processed is important.

Driven by the unprecedented volumes of data from experiments, measurements, and large-scale simulations, we are dealt with a vast amount of data. Here is where we expect ML to play its role and solve the problems in a more efficient and less time-consuming manner. ML is applicable to tackle many problems in fluid mechanics such as reduced-order modelling (ROM), optimization and control. In this work, we limit our focus to only one sub-domain of fluid mechanics known as the multiphase flow systems. We will model the liquid jet in crossflow (LJIC) systems (Chapters three and four), and categorize them into sub-flow regions using ML techniques. A brief overview of LJIC systems and different ML frameworks will be discussed in the next few pages.

1.1 Liquid Jet in Crossflow (LJIC)

The liquid jets in crossflow or transverse jets have become widely used in a variety of technological applications, including fuel injection in gas turbine engines, rocket vehicles, exhaust plumes from power plants, etc. (Wu et al. (1997)). Liquid jets are typically injected from nozzles or the walls of combustors into the airstream under crossflow conditions. A schematic of the procedure could be found in Fig.1.1.

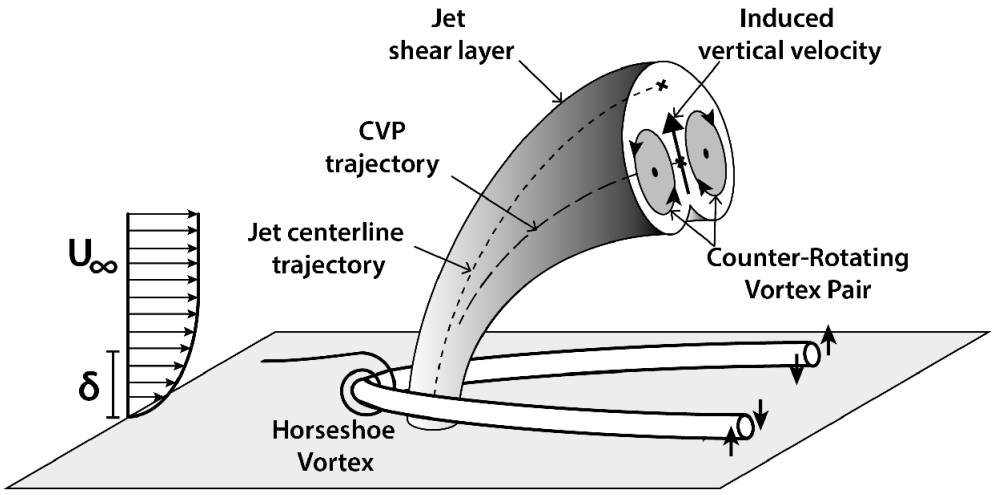


Figure 1.1: Global flow field of a jet in crossflow, in which a counter-rotating vortex pair (CVP) is formed (Cambonie et al. (2013)).

It may be directly understood from the figure that despite the possibility of the jet injection into the crossflow at a steady rate, the unsteady behaviour of the evolution of flow structures is vividly evident. The periodic nearfield rollup of vortical structures contributes to the creation of vortices that dominate the ensemble-averaged flowfield principally the counter-rotating vortex pair (CVP) structure observed to dominate the transverse jet's cross-section in the mean. In addition to the CVP, there are jet shear layer vortices formed in the nearfield, horseshoe vortices forming in the plane of the injection wall and the upright wake vortices that are generated inside the wall boundary layer and then shed beyond to the jet.

Despite the apparent simplicity of this phenomenon, the modelling of such processes has been the subject of many experimental and theoretical researches.

Extensive knowledge of the exploration of LJIC and their applications may be found in (Margason (1993)). It is well believed that the instability waves can develop at the interface of the gas and liquid phases and grow further until the jet disintegrates into large ligaments which will go through further breakup themselves.

Breakup Regime Map of LJIC

A good understanding of LJIC would be achieved in the identification process of the different flow behaviours denoted as the breakup regimes that occur at specific flow conditions. These conditions are not universal and different opinions have been mentioned by researchers in the literature. One of the most well accepted of them is the work of Wu et al. (1997), where they characterized the regime map of the LJIC into two sub-regimes (column and surface breakup). This classification was based on the jet-to-air momentum flux ratio (q) and the gaseous Weber number (We). In the column breakup regime, waves grow on the jet surface that leads to column fracture without remarkable mass shedding from the column surface. On the other hand, the stripping of the ligaments and the droplets from the surface of the jet is known as the surface breakup regime that occurs at high jet-air momentum flux ratios and Weber numbers. In some applications, both regimes may be present given the fact that the dominance of each regime over the other would be of significant importance in there.

Wu et al. (1997) also divided the column breakup section into four sub-categories according to the rise of the Weber number but without any dependence to the q. These categories are: enhanced capillary, bag, multimode and the shear breakup regimes. Similar breakup maps were also depicted in the works of Mazallon et al. (1999) and Sallam et al. (2004). Sallam et al. (2006) categorized the breakup of a turbulent transverse flow into two major classes of aerodynamic and turbulent breakup regimes. These regimes were separated using a non-dimensional parameter ($(We)q^{1/3}$). It is clear that breakup regimes map is not unique and is a function of the chosen non-dimensional parameter. Madabhushi et al. (2006) incorporated the effect of Reynolds number and suggested the partnership of the Weber number with the Reynolds number for mapping the atomization process.

Reynolds number is linked to We and q through the relation below:

$$Re_l^2 = qWe/Oh^2 \tag{1}$$

In this relation, Ohnesorge number ($Oh = \mu_l/(\rho_l d\sigma)^{0.5}$) groups the parameters related to atomization. Re_l is, however, a direct estimate of whether the jet at the orifice is laminar, fully turbulent, or is transitioning to turbulence (turbulent core and laminar surface), and therefore it is more relevant than q as a global parameter for determining the atomization regime (Madabhushi et al. (2006)). Throughout this study, we will use the regime map suggested by Wu et al. (1997). The map is visualized in Fig.1.2.

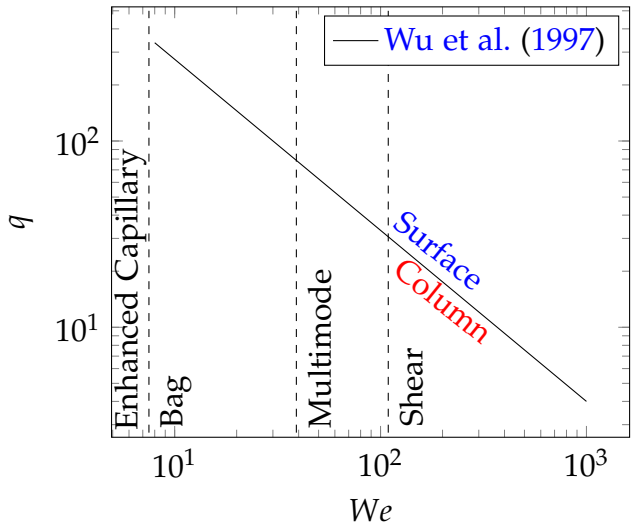


Figure 1.2: $We - q$ regime map of primary breakup processes of a nonturbulent liquid jet in a crossflow. Reprinted from Brown et al. (2006).

1.2 Reduced Order Modelling (ROM)

Analysis of dynamical systems by the extraction of spatio-temporal patterns from the data generated either by numerical simulation or physical experiment has been widely used in a vast range of fields as fluids, physics, and machine learning (Fujii and Kawahara (2019)). Dealing with big data information in multiphase flow and thermal spray applications, it is of considerable significance to take advantage of the modal decomposition methods such as proper orthogonal

decomposition (POD) and dynamic mode decomposition (DMD) to extract dynamic information from the flow fields. In both frameworks, the data is usually in the form of modes, each of which is representative of a specific flow regime. The physical mechanisms embedded in the modes are on a dynamical system of fewer degrees of freedom compared to the initially given system (Schmid (2010)). The distinction between the two methods is that POD determines the optimal set of modes to represent the data based on the energy norm, while DMD captures dynamic modes with associated growth rates and frequencies.

Of particular obstacles in reaching an optimal spectrum, is the existence of noising in the system. Adding more modes does not necessarily increase the precision of the decomposition procedure while it can bring extra noise into the system. Therefore, taking an optimum rank truncation number seems crucial as well as other options like data centering, adding penalty function, etc. (Dang et al. (2018); Hirsh et al. (2019)). Another challenge while working with the conventional ROM frameworks is their limitation in dealing with high Reynolds number flows. With POD, for example, 7260 modes are required to reproduce 95% of the total energy for a turbulent channel flow at Re=180, while we need only 2 POD modes to cover 99% of the total energy for a flow around a circular cylinder at Re=100 (Proctor et al. (2016)). In recent years, machine learning has been also widely applied in the field of fluid dynamics and is well-known for its power to account for nonlinearity (Fujii and Kawahara (2019)). This task is performed usually through neural networks of multiple layers. However, we will cover the application of two classic machine learning algorithms of SVM and RF in the last chapter of this thesis. In continue, a brief history of the previous works on the reduced-order modeling of different liquid jet systems is reported.

Early descriptive works in POD goes back to 1993, where Berkooz et al. (1993) took advantage of it in the analysis and modeling of turbulent flows although it was already a well-established tool for statistical data analysis and data compression. The claim that a physical understanding of the underlying system is required besides the promising mathematical approaches, was mentioned at that time and is still considerable in research. Recent works have been focused on atomization and liquid jet breakup where POD is suitable for understanding underlying dynamical processes that control the system. For instance, Arndt et al.

(1997) applied POD to the far field of an axisymmetric turbulent jet and reported the dominant frequency peaks. It was the first time that POD was applied to the pressure field surrounding the jet. Ease of measurement of the pressure fields and the resulting simplified data analysis set the stage for POD-based control strategies. POD analysis was applied to investigate the jet fluctuation appearing in the far field for the standard case of large velocity ratios. It was also directed to the study of the transition between the stable and unstable flow conditions around the critical velocity ratio (Bernero and Fiedler (2000)). Both velocity and vorticity fields were employed to perform this study where the velocity-based POD appeared to produce better results than the vorticity POD.

The circulation zone of an annular jet was addressed by particle image velocimetry (PIV) and POD was applied to find the relationship of the radial frequencies to the inner structures of the instantaneous PIV fields. This statistical method could help to find the overall behaviour of the flow and to link the fluctuations with typical modes. It was shown that the flow could be decomposed into four main modes where each is responsible for a characteristic motion of the circulation zone (Patte-Rouland et al. (2001)). Dynamics of flame was revealed as a set of statistical quantities referred as modes. The flame fluctuations are used to derive empirical functions representing the most important features of the flame while the validation of the capabilities of the techniques is limited to the case of an unsteady laminar flame (Duwig and Iudiciani (2010)).

The impact of the density ratio on the liquid core dynamics of a turbulent liquid jet injected into gaseous crossflow was studied by Herrmann et al. (2011). They engaged POD to extract the wavelength of the most dominant feature associated with a travelling wave along the jet where it was shown that the increase in density ratio results in a decrease in the amounts of wavelength. The decrease in wavelength was further correlated to the noticeable increase in liquid core penetration with reduced bending in the crossflow and spreading in the transeverse directions.

The primitive work on DMD was done by Schmid (2010), where a new methodology capable of extracting dynamic information from the flow fields was introduced. Formerly, the Arnoldi method and various other decomposition techniques were available to provide the tools to extract the dynamics of the flow

patterns of the simulation data that are more instructive in describing the fluid process with respect to POD. However, the newly offered DMD technique was the first to provide temporal dynamic characteristics together with spatially coherent structures for the experimental data. The modern DMD technique was once again highlighted in the work of Schmid (2011) and demonstrated on image-based flow visualization. The technique was applied to PIV data and image-based flow visualizations where the data collection was performed on a numerical simulation of a flame based on a variable-density jet and on experimental data from a laminar axisymmetric water jet. It was demonstrated that DMD has vivid advantages over POD as the former struggles for a representation of the dominant flow features within a temporally orthogonal framework (pure frequencies), while the latter is based on a spatially orthogonal model. A general understanding of different decomposition techniques could be achieved in the review paper of Taira et al. (2017).

An experimental investigation of the near field pressure of a compressible jet was performed in the work of Mancinelli et al. (2018) where the POD modes were interpreted in terms of their hydrodynamic nature through the computation of the frequency spectrum of the reconstructed space-time pressure fields using each mode. Higham et al. (2018) showed that the nonlinear dynamics can arise in the transition regimes of shallow flows to a quasi-2D behaviour which can result in identifying structures composed of multi frequencies that is a sign of convoluted dynamics. Thus DMD was recommended for the analysis of such phenomena.

Prakash et al. (2018) attempted to study the liquid jet breakup parameters with respect to the air flow and liquid jet entry conditions. It was observed that the dependence of the trajectory of the spray is not just limited to the momentum flux ratio, q, but also requires correction factors with respect to the injection entry conditions, which are in turn related to L/D values. L/D parameter was varied between 10 and 100 in order to obtain fully-developed laminar flow, transition and turbulent flow. Transient analysis of the liquid jet breakup phenomena was conducted by subjecting the high-speed images to POD analysis where the liquid jet manifested various different modes of breakup at different oscillating frequencies. While both the laminar and turbulent jets exhibited whiplash-like action, the

turbulent jet was observed to be more vigorous showing higher oscillation amplitudes and shorter wavelengths for the corresponding modes. It was correlated to the great tendency of laminar jets to penetrate much further than the turbulent jet. Recently, Murata et al. (2020) have developed a novel nonlinear mode decomposition technique to visualize the decomposed flow fields. Convolutional neural network auto-encoders (CNN-AE) have been employed in their work. CNNs are a class of deep neural networks mostly applied to analyze visual imagery.

The architecture of a CNN consists of an input and an output layer as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other dot product. The activation function is commonly a rectified linear unit (ReLU) layer, and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers the details of which could be accessed in the work of LeCun et al. (1999).

The activation layer of the CNN network in the work of Murata et al. (2020) was modified over different linear and nonlinear functions to analyse the resulting models with the ones extracted from POD and it was shown that the decomposed fields are similar to those of POD in case of a linear activation layer while with the nonlinear activation function, L_2 norm error of the reconstruction was reduced as compared to those of POD. This was a breakthrough in the data-driven discovery of fluidic systems due to the capability of the CNNs in representing more information with the same number of modes. However, the performance of the deep learning algorithms is highly controlled by the amount of data. Therefore, the required data collection could be overwhelming while dealing with experiments. In this case, a conservative approach seems logical in which we may be victorious in exploiting the power of machine learning algorithms while we own a limited amount of data. This issue will be further discussed in the fourth chapter. However, a general idea of different machine learning algorithms and the ones that we will use throughout this research are represented in the next section.

1.3 Machine Learning (ML)

Fluid mechanics has traditionally dealt with massive amounts of data from experiments, field measurements, and large-scale numerical simulations (Brunton et al. (2020)). Indeed, in the past few decades, big data have been a reality in fluid mechanics research (Pollard et al. (2016)) due to high-performance computing architectures and advances in experimental measurement capabilities. Over the past fifty years, many techniques have been developed to handle such data, ranging from advanced algorithms for data processing and compression to fluid mechanics databases (Perlman et al. (2007)). There are several factors which are each game changers in the sense that they encourage more focus on data-driven problem solving in the future. These include the persistent increasing trend of data volume, the abundance of open-source software for computation, and etc. These levels of evolution in different fields have pumped a huge amount of investment in the field of machine learning (ML). ML is making its role as a leader in different industries, specifically fluid mechanics. ML provides us with a wide range of algorithms that could each be taken advantage of to cope with challenges in fluid mechanics, such as reduced order modelling (ROM), turbulence modelling, and flow control. These learning algorithms may be categorized into supervised, semi-supervised, and unsupervised learning (Fig.1.3).

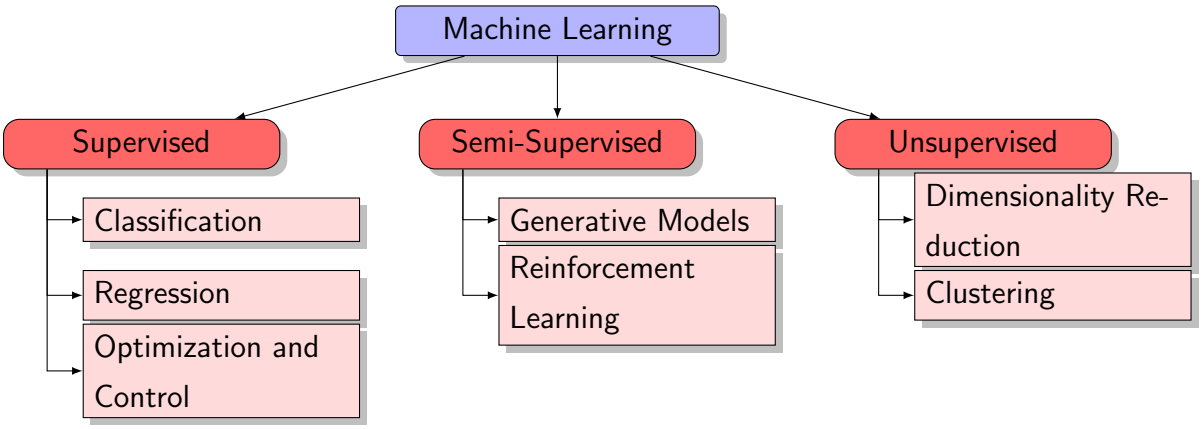


Figure 1.3: Categorization of machine learning algorithms into supervised, unsupervised, and semi-supervised. Adapted from Brunton et al. (2020).

Throughout the current study, we will investigate two applications of machine learning into our case study (LJIC) which was introduced in the previous chapter. Dimensionality reduction techniques will be deployed as unsupervised ML algorithms to model the physics of our case on a reduced order basis. These techniques that are called the reduced order models are extensively discussed in the next section. Additionally, we will use the classification capabilities of ML for categorizing the breakup regimes of the LJIC cases relying on the reduced order models that we will generate as the features.

1.4 Objectives

The main motivation for the current study is to address the potential merit of data-driven modelling techniques to obtain explainable solutions for a number of benchmark and complex problems in fluid mechanics specifically in the field of multiphase flow.

The objectives are summarized below:

1. Investigation of the spatial and temporal governing modes of the liquid jets in crossflow.

2. Extracting the dominant frequencies of the jet oscillations and studying the effects of them in the flow characteristics.

3. Studying the effect of the nozzle aspect ratio in analyzing the liquid jets in crossflow systems using dimensionality reduction methods.

4. Classifying the flow regimes of the liquid jets in crossflow using the combination of supervised and unsupervised machine learning algorithms

Chapter 2

Methodology

2.1 Proper Orthogonal Decomposition (POD)

A short review of the POD technique and its use in deriving reduced models for the systems is provided here. Dynamical systems consisting of temporal and spatial processes can be represented by considering the PDE systems of the form:

$$U_t = N(U, U_x, U_{xx}, ..., x, t) \tag{2}$$

where U is a vector of quantities, and the subscripts t and x denote partial differentiation. The function $N(.)$ captures the spatio-temporal dynamics that are specific to the considered system (Chatterjee (2000)).

By the consideration of the variable and basis expansion technique, the standard eigenfunction expansion assumes a solution of the form:

$$u(x,t) = \sum_{n=1}^{\infty} a_n(t)\phi_n(x) \tag{3}$$

where $\phi_n(x)$ can be any orthogonal set of functions such that:

$$(\phi_j(x), \phi_k(x)) = \delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{4}$$

where the δ_{jk} denotes the Dirac delta function. In general, it is common to use a set of eigenfunctions that produce a rapid evaluation of the solutions of (2), i.e., the fast Fourier transform (FFT). However, the optimal POD basis functions

decomposed from a singular value decomposition (SVD) of the representative dynamics of the system are the best low-dimensional framework. A general formulation of SVD that is a built-in subroutine in many software environments as MATLAB or Python Numpy library comes in continue.

In order to build a complete set of POD modes, a data set represented as the matrix X is required. Each column of the matrix is a snapshot that is taken at evenly spaced values in time. Therefore, if the data consists of m spatial points at n snapshots, then $X \in \ C^{m\times n}$. The SVD breaks up the matrix X into three matrices:

$$X = U\Sigma V^* \tag{5}$$

where $U \in \ C^{m\times n}$, $V \in \ C^{m\times n}$ and $\Sigma \in \mathbb{R}^{m\times n}$ (real numbers) and the asterisk denotes the conjugate transpose. The data matrix X and its SVD factorization (5) can be expressed as below (6)-(7):

$$X = \begin{bmatrix} x_1^1 & x_2^1 & . & . & x_n^1 \\ x_1^2 & x_2^2 & . & . & x_n^2 \\ . & . & . & . & . \\ . & . & . & . & . \\ x_1^m & x_2^m & . & . & x_n^m \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix} \tag{6}$$

$$SVD(X) = \begin{bmatrix} \overrightarrow{\phi_1} \\ . \\ . \\ . \\ \overrightarrow{\phi_i} \\ . \\ . \\ . \\ \overrightarrow{\phi_m}. \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & . & . & . & . & . & 0 & 0 \\ 0 & . & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & \sigma_j & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & . & 0 \\ 0 & 0 & . & . & . & . & . & 0 & \sigma_n \end{bmatrix} \begin{bmatrix} \overrightarrow{\psi_1} \\ . \\ . \\ . \\ \overrightarrow{\psi_i} \\ . \\ . \\ . \\ \overrightarrow{\psi_m} \end{bmatrix} \tag{7}$$

.

The matrix Σ is a diagonal matrix with non-negative elements σ_j. The (σ_j)s are referred to as the singular values of X and are ordered such that:

$$\sigma_1 \succeq \sigma_2 \succeq ... \succeq 0.$$

the matrices U and V are composed of the eigenvectors $\overrightarrow{\phi_i}$ (rows of U) and $\overrightarrow{\psi_i}$ (rows of V^*) of the covariance matrices XX^* and X^*X, respectively. As a result, SVD returns a complete set of basis functions for the columns of the data matrix (Chatterjee (2000)).

One way to choose an optimal subset of POD modes is to determine the cumulative energy of that subset defined as:

$$E_c = \frac{\sum_{i=1}^{r} \sigma_i^2}{\sum_{i=1}^{n} \sigma_i^2} \tag{8}$$

The dimensionality reduction could also be performed using the relative energy E_r of one specific (jth mode) defined as:

$$E_r = \frac{\sigma_j^2}{\sum_{i=1}^{n} \sigma_i^2} \tag{9}$$

where the total energy is normalized as $\sum_{i=1}^{n} \sigma_i^2 = 1$ (Chatterjee (2000)) It is usual that a truncation scheme where the energies of the truncated POD modes tend to reach 99% of the total energy is a desirable one.

2.2 Dynamic Mode Decomposition (DMD)

A short review of the DMD technique and its use in deriving reduced models for the systems is provided here. In the DMD scheme, we usually collect the data from a dynamical system in the following form (Kutz et al. (2016)):

$$dx/dt = f(x, t; \mu) \tag{10}$$

where $x(t) \in \mathbb{R}^n$ is the representative vector of the state of the system at time t, μ includes the system parameters, and *f(.)* specifies the dynamics. Converting the continuous-time dynamics from (10) will result in the discrete-time representation of the form:

$$x_{k+1} = F(x_k) \tag{11}$$

The DMD scheme takes the system measurements alone to approximate the constructing dynamics. The DMD approximates the dynamical system (10) in a linear system of the form:

$$\frac{dx}{dt} = Ax \tag{12}$$

That has a well-known solution of the form below, given the initial condition x(0) (Tu et al. (2013)):

$$x(t) = \sum_{k=1}^{n} \varphi_k exp(\omega_k t) b_k = \phi exp(\Omega t) b \tag{13}$$

where φ_k and ω_k are the eigenvectors (DMD modes) and eigenvalues of the matrix A, and the coefficients b_k are the coordinates of x(0) in the eigenvector basis. The continuous-time map matrix A can also be replaced by the discrete-time map $\mathbb{A}$ as below:

$$\mathbb{A} = exp(A\Delta t) \tag{14}$$

Therefore, the solution may be easily expressed in the discrete matrix form using the eigenvalues λ_k and eigenvectors φ_k of the discrete-time map $\mathbb{A}$ (Tu et al. (2013)):

$$x_{k+1} = \mathbb{A} x_k \Longrightarrow x_k = \sum_{k=1}^{n} \varphi_j \lambda_j^k b_j = \phi \Lambda^k b. \tag{15}$$

The DMD algorithm approximates the measured trajectory x_k for (k = 1,2,...,m) in a least-square sense so that the following error function is minimized across the whole spatial domain:

$$\parallel x_{k+1} - \mathbb{A} x_k \parallel_2 \tag{16}$$

To aim this purpose, the total number of snapshots (n) will be split into two large datasets:

$$X = [x_1 \mid x_2 \mid ... \mid ... \mid x_{n-1}], X' = [x_2 \mid x_3 \mid ... \mid ... \mid x_n]. \tag{17}$$

The locally linear approximation can be rewritten in terms of these datasets:

$$X' \simeq \mathbb{A} X \tag{18}$$

The optimum $\mathbb{A}$ matrix is given by:

$$\mathbb{A} = X' X^{\dagger} \tag{19}$$

where † is the Moore-Penrose pseudoinverse [1]. To simplify the process, we project our data onto a lower rank subspace by using (n-1) POD modes and then work with the lower rank mapping operator in order to reconstruct the eigenvalues

[1] A general form of the inverse operator supporting matrices with real and/or imaginary elements

and eigenvectors of operator $\widetilde{A}$ without even the need to compute $\widetilde{A}$ directly. The DMD algorithm may be continued as follows (Tu et al. (2013)):

1. Singular value decomposition of X (Trefethen and Bau (1997)):

$$X \simeq SVD(X) \Longrightarrow X \simeq U\Sigma V^* \tag{20}$$

where $U \in \mathbb{C}^{m\times r}$, $V \in \mathbb{C}^{n\times r}$ and $\Sigma \in \mathbb{R}^{r\times r}$ and r is the rank of the reduced SVD of the X matrix. This step is done in order to find a low-dimensional structure in the dataset that is possible by calculating the relative energies of the modes defined in (8) decreasing sharply to zero after a few number of dominant modes.

2. The matrix $\mathbb{A}$ is found by multiplying the pseudoinverse of X and the conjugate transpose of U:

$$\mathbb{A} = X'V\Sigma^{-1}U^* \tag{21}$$

More computationally efficient is to compute $\tilde{\mathbb{A}}$, the projection of the full matrix $\mathbb{A}$ onto POD modes:

$$\tilde{\mathbb{A}} = U^*\mathbb{A}U = U^*X'V\Sigma^{-1} \tag{22}$$

3. Eigen-decomposition of $\tilde{\mathbb{A}}$:

$$\tilde{\mathbb{A}}W = W\Lambda \tag{23}$$

Where the columns of W are eigenvectors and Λ is a diagonal matrix containing the corresponding eigenvalues λ_k.

4. Reconstruction of the eigen-decomposition of $\mathbb{A}$: eigenvectors of $\mathbb{A}$ can be easily calculated using the eigenvectors of $\tilde{\mathbb{A}}$ and the eigenvalues Λ which is common between $\mathbb{A}$ and $\tilde{\mathbb{A}}$.

$$\phi = X'V\Sigma^{-1}W \tag{24}$$

The formulation used above is the so-called exact DMD method which is different than the version proved by Tu and Rowley (2014) as the projected DMD.

A brief definition of the two proposed methods is given in Table.2.1.

Table 2.1: Summary of the modal decomposition techniques for fluid flows in this thesis

Methods	Descriptions
POD	Finds an optimal number of modes to truncate the data based on the modes' energy content (L2 norm).
DMD	Extracts the dynamic underlying patterns with the corresponding growth rates and frequencies.

2.3 Method Demonstration

To demonstrate the efficacy of the POD and DMD algorithms in decomposing the input signal into its components, two case studies are considered:

1) A mix signal of two spatiotemporal waves

2) A Re = 100 flow around a cylinder wake which is approximately similar to the near-field results of the liquid jet in crossflow.

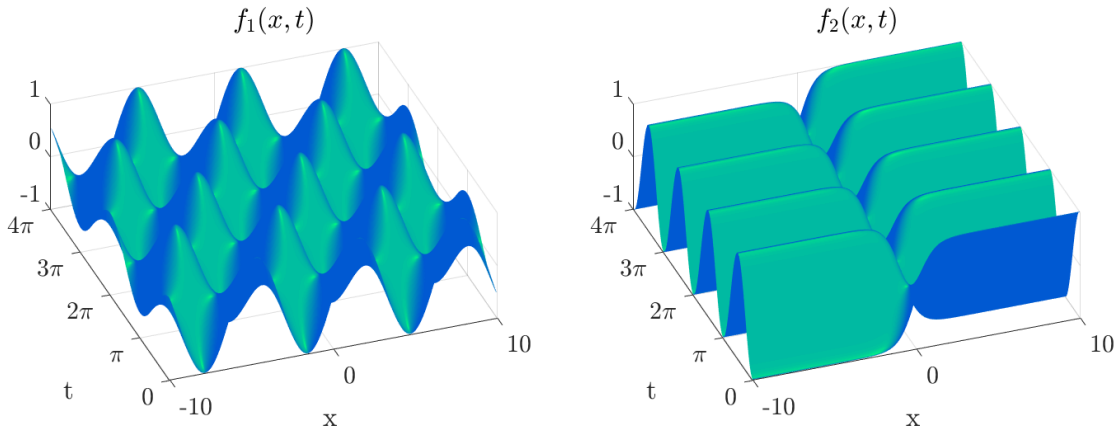
2.3.1 Mixed Signal of Two Spatiotemporal Waves

In this example, we consider a simple mixed spatio-temporal signal and the goal is to analyze the performance of POD and DMD algorithms to decompose the signal into its parts. The signal under study is:

$$f(x,t) = f_1(x,t) + f_2(x,t) = sin(x)exp(i1t) + tanh(x)exp(i2t) \tag{25}$$

These two spatiotemporal signals $f_1(x,t)$ and $f_2(x,t)$ are generated using 400 data points evolving in 200 time frames and then depicted in Fig.2.1.[2] Two pertaining governing frequencies embedded in the physics of this wave are ω_1=1 and ω_2=2. We expect that the reduced order modelling techniques (POD & DMD) can easily extract the base modes and explore out the true frequencies. Their produced modes could be then utilized to rebuild the entire mixed signal which is displayed in Fig.2.2.

[2]The Python code is available in Appendix.A to construct the spatiotemporal signals.



(a) (b)

Figure 2.1: Spatio-temporal dynamics of each of the two waves, a)$f_1(x,t)$, b)$f_2(x,t)$

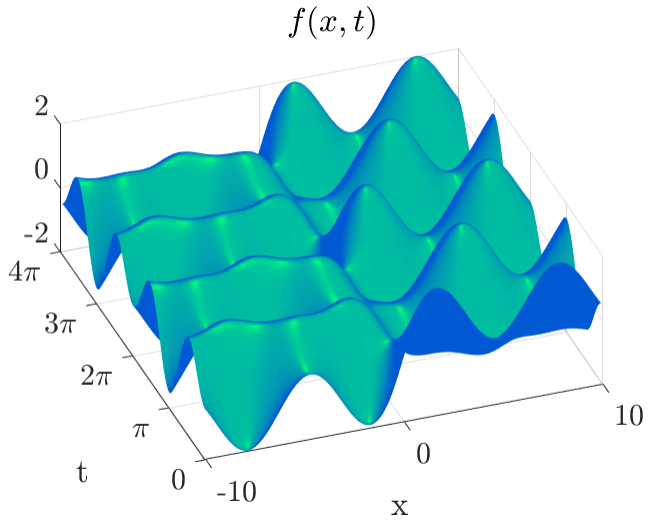


Figure 2.2: Mixed signal $f(x,t)$

Results indicate that while both methods can capture the governing physics in two modes, POD is incapable of truly finding the exact sine and hyperbolic tangent behaviours but produces two underlying waves each of which is a mixture of true modes. On the other hand, DMD modes closely matches with the hidden components of $f_1(x,t)$ & $f_2(x,t)$. It can be shown by spatial and temporal spectra in Figs.(2.3-2.4)

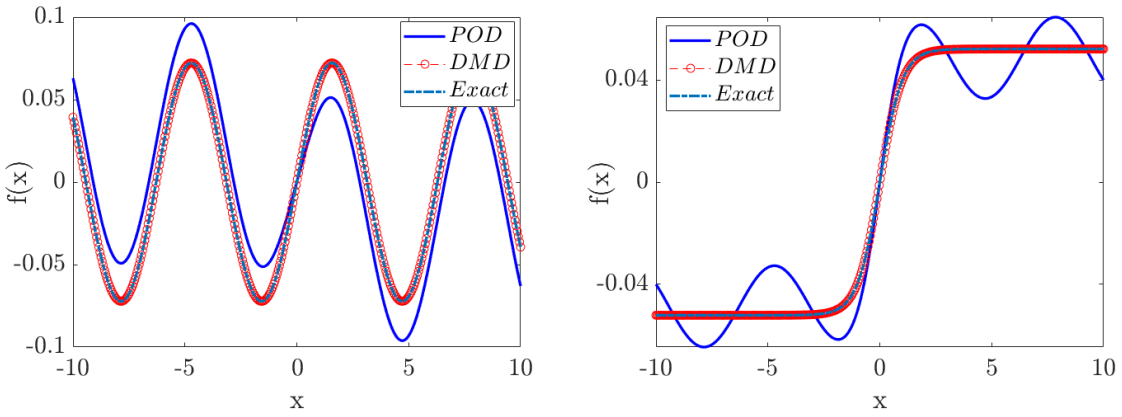


Figure 2.3: Spatial spectra of the mixed signal, a) mode 1, b) mode 2

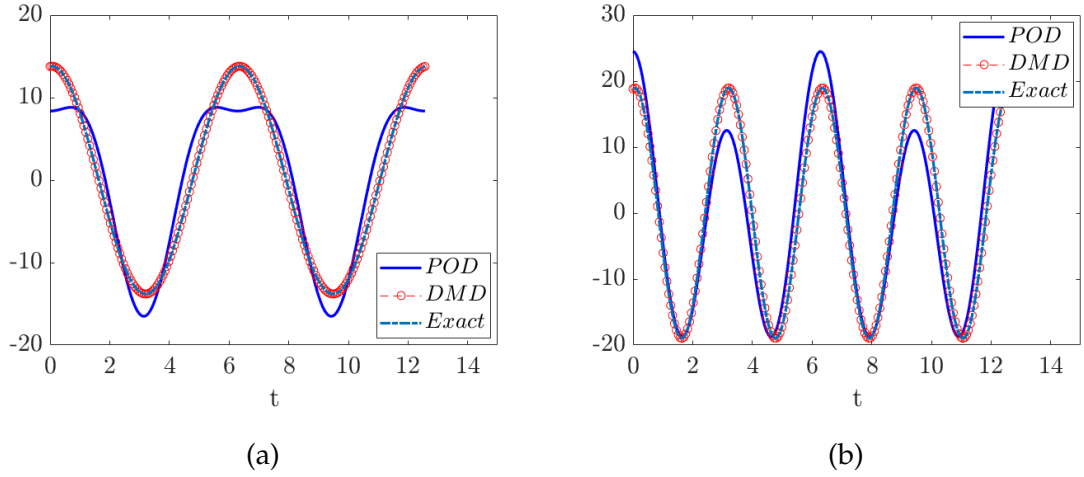


Figure 2.4: Temporal spectra of the mixed signal, a) mode 1, b) mode 2

It is important in the sense that POD tries to decompose the mixed signal into orthogonal components that are chosen to maximize the variance in the data but not necessarily able to identify the independent parts like DMD (Kutz et al. (2016)). The error pertaining to the POD method is as small as only 1.5538×10^{-4} and could be calculated using the mean-squared error[3]. This amount is approximately equal to 0 for the DMD method. Irrespective of their accuracy, both methods truncate the initial rank-200 data onto a correct rank-2 structure.

[3]Mean squared error is defined in Appendix and is found using the MATLAB imsse function.

2.3.2 Von Karman Vortex Street Past A Cylinder

Problem

A Karman vortex street is known as a recurring pattern of swirling vortices, made by the separation of flow of a fluid around blunt bodies typically called the vortex shedding process. A street of vortices will be shaped usually over a critical Reynolds number of 90. The global Reynolds number for a flow is a measure of the ratio of inertial to viscous forces in the flow of a fluid around bodies or inside channels. It may be defined as the non-dimensional global speed parameter of the whole fluid flow.

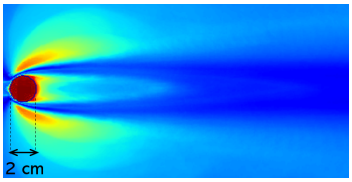
$$Re_l = ul/\nu \tag{26}$$

where u is the free stream flow speed, l remarks the characteristic length parameter which is the diameter of the cylinder here and ν denotes the free stream kinematic viscosity.

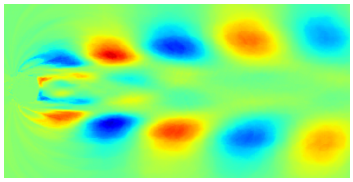
Simulation and Results

For simulating the above described problem, the ICEM CFD software has been used for generating the mesh and the boundary conditions around a cylinder with the total diameter of 2cm. Afterwards, the two-dimensional Navier-Stokes equations are solved in ANSYS Fluent based on the multidomain method of (Colonius and Taira (2008)). A time step size of $\Delta t = 0.01s$ has been also chosen to satisfy the CFL condition. Contours of velocity are produced and then collected at specified intervals in time, $20\Delta t$, leading to a sampling frequency of five frames per second.

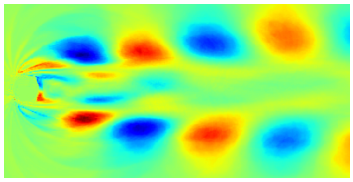
To model the fluid flow dynamics on a reduced order system, each of the velocity snapshots are reshaped into a tall skinny vector x_k, and the vectors constitute the columns of the matrices X and X', as described in (17). Afterwards, SVD-based techniques of POD and DMD will be implemented and the spatial modes are generated in result (Figs.2.5a-2.5e).



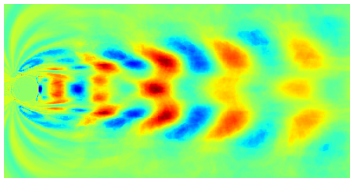
(a)



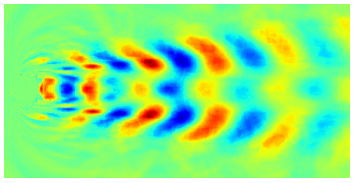
(b)



(c)



(d)



(e)

Figure 2.5: Spatial domain of the von Karman street past a cylinder, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4

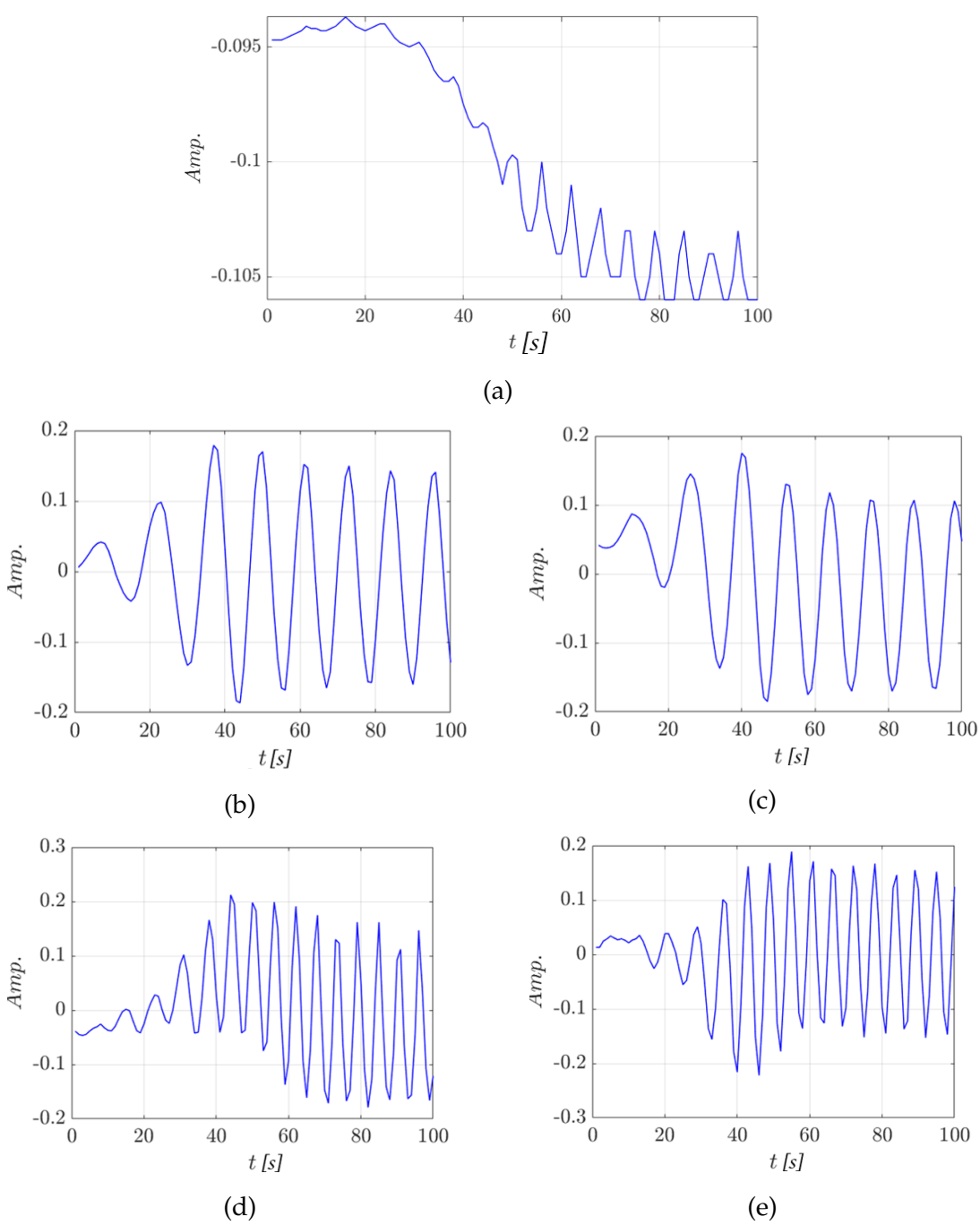


Figure 2.6: Temporal spectra of the von Karman street past a cylinder, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4

As it can be seen, Mode 0 (Fig.2.5a) shows the average flow that portrays the flow separation on the cylinder while modes 1 and 2 (Fig.2.5b-2.5c) accurately capture the von Karman vortex street behaviour. On the other hand, Modes 3 and

4 (Fig.2.5d-2.5e) are representing a convection physics past the cylinder towards the downstream. These spatial structures are characterized with positive and negative pixel values shown with an *hsv*-type colormap using matplotlib library in Python where for a given mode, each of red (positive) or blue (negative) pixels determines the probability of the existence of the fluid in those locations. It means that, for a single mode, if the liquid is present where the values are positive, then it is unlikely to be present where the values are negative. The modal frequency analysis could also tell us how frequently the liquid is present where the values are positive or negative.

It can be understood from (Figs.2.6a-2.6e) that modes 1,2 and modes 3,4 are paired modes with 90° phase shift with respect to each other. This fact is easily comprehensible by plotting the temporal dynamics of the paired modes versus each other. For example, the temporal terms of modes 1 and 2 are plotted versus each other in (Figs.2.7) where spin-offs around a circle with a known diameter is obvious. As we may observe, this spinning behaviour stabilizes around a circle of specified diameter which is due to the development of the von Karman vortices from the first snapshot all the way to the steady state condition.

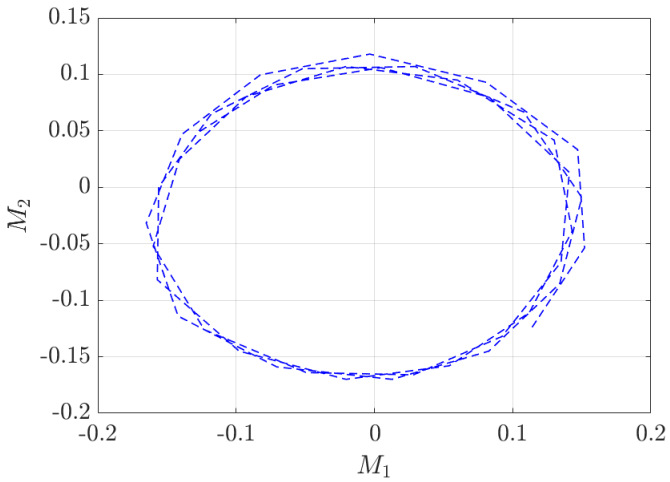
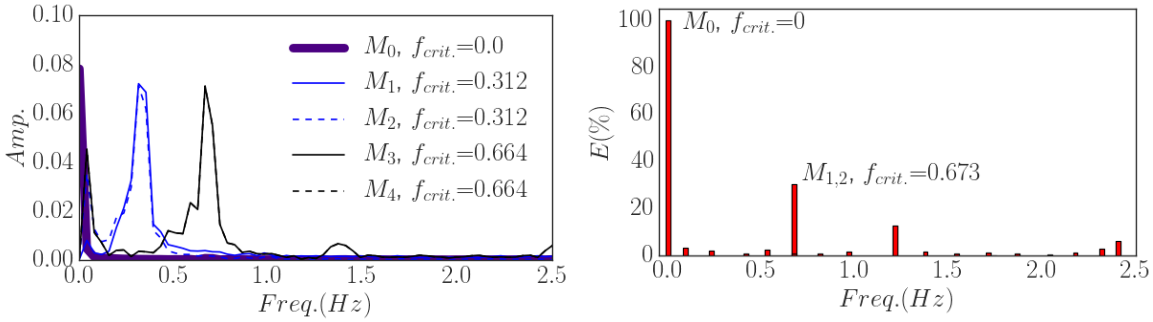


Figure 2.7: Temporal spectra of the von Karman street past a cylinder, mode 1 vs. mode 2

The modal frequency analysis could be discussed for POD and DMD schemes

to find the corresponding Strouhal numbers. As POD is solely incapable of identifying the underlying frequencies, a power spectral density (PSD) diagram will be produced by transforming the POD modes' temporal terms (V vectors) onto the Fourier space. Plotting the PSD versus the frequency spectrum, we can find the critical frequencies where the maximal power spectral densities will happen (Fig.2.8a).



(a) (b)

Figure 2.8: Modal frequency spectrum for the flow around cylinder at Re=100 a) POD-PSD, b) DMD energy

We can now extract the frequencies from POD and DMD and compare them with the experimental results using the dimensionless number called the Strouhal number (St) defined as below:

$$St = \frac{f_s D}{u_\infty} \tag{27}$$

Where f_s is the shedding frequency, D is the cylinder diameter (2 cm) and u_∞ is the free-stream velocity (0.0723 m/s).

As it is evident, DMD extracts a single frequency value equal to 0.673 which leads to an Strouhal number of 0.186 that is well verified with the experimental results. PSD spectrum extracted from the POD temporal terms also peaks at the critical frequency of 0.664 that gives an Strouhal number of 0.183 that is again satisfying the experimental results.

It is notable that we could use either the vorticity or the velocity fields for the data-driven modelling. Velocity data snapshots were used here to generate the reduced model.

2.4 Hierarchical Investigation of Liquid Breakup Systems

In this section, we apply the methods introduced in the previous section to the datasets of liquid injection systems to better understand the fundamental physics behind the basic fluid flows which will then enable us to investigate more practical systems like the liquid jets in crossflow. The emphasis will be provided on creating an insight into the interpretation of modes in a hierarchical structure due to the lack of a strong link between the decomposition theory and the physical meaning of the modes.

We will demonstrate again that the modes which capture the true physical phenomena produce subsequent modes that occur at higher harmonics up until the Nyquist limit. This point was also highlighted in the first chapter where we analyzed the decrease of the modal structure scales by gradually increasing the modal frequency.

2.4.1 Laminar Jet

case 1: $Re \approx$ **709,** $We_l \approx$ **16** [4]

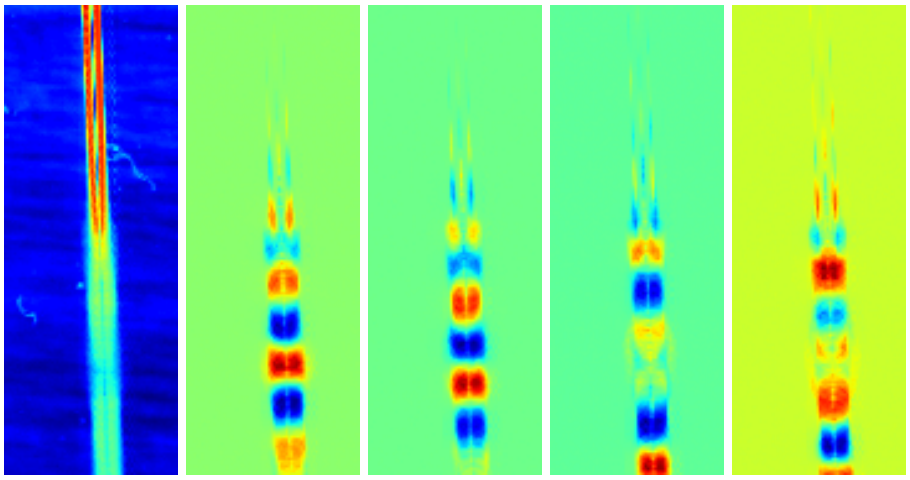
The most naive case to be surveyed for POD and DMD in the jet breakup systems is the laminar jet where the image snapshots have no significant variance during the breakup's course of action. Hence, the presumption of the existence of a single particular mode seems rational.

The first five POD modes are reshaped and illustrated in Fig.2.10 along one data snapshot (Fig.2.9)[5].

[4] **The same setup as the one described in the section 4.2.1 is used here; the liquid is water, orifice diameter** D **= 0.43 mm and the inlet velocity** u **= 1.65 m/s.**

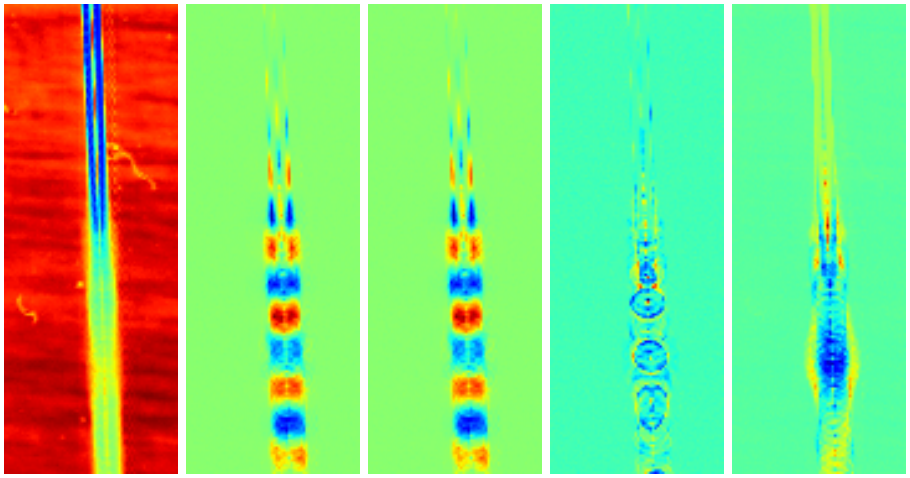
[5] Background elimination has been applied on all snapshots using the ImageJ software

Figure 2.9: Capillary breakup of a laminar jet



(a) (b) (c) (d) (e)

Figure 2.10: First 5 POD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the laminar jet ($Re \approx 709$, $We_l \approx 16$)

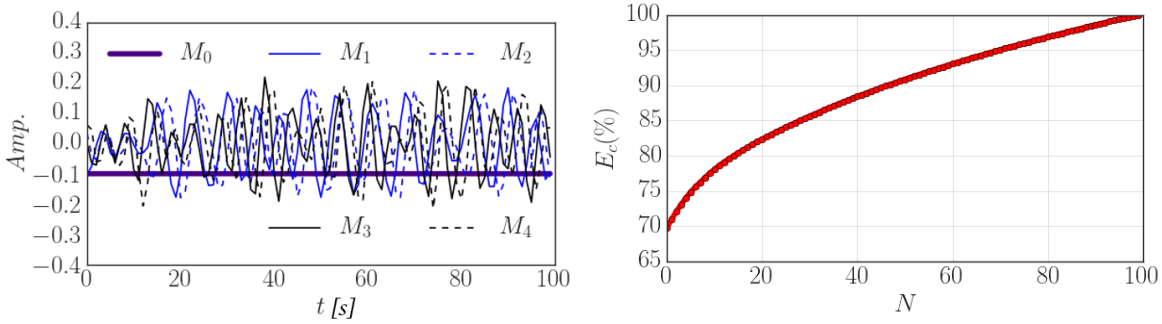


(a) (b) (c) (d) (e)

Figure 2.11: First 5 DMD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the laminar jet ($Re \approx 709$, $We_l \approx 16$)

First conclusion to be made is the existence of the mode 0-*th* in both POD and DMD cases. This mode displays the average of all image frames that is characterized with a frequency value of 0 Hz.

It is also visible that the pixel intensity values are inverted from POD to DMD. It is noticeable that this has no great effect on comprehending the physics of the problem and the first POD/DMD modes are equivalent.



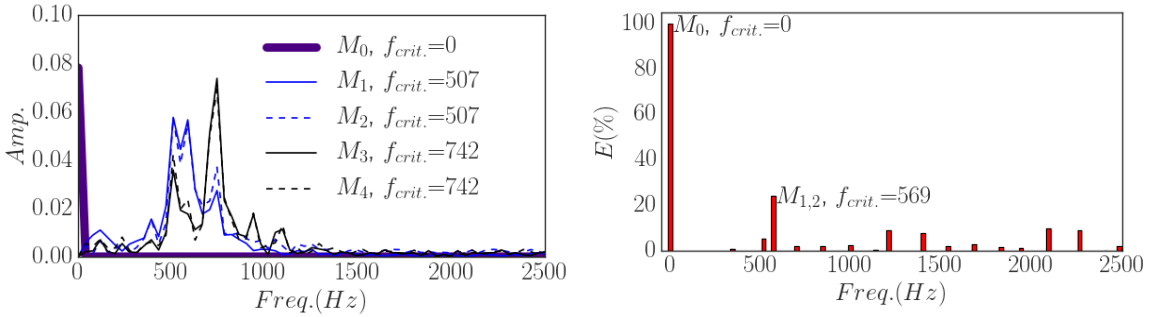
(a) (b)

Figure 2.12: a) temporal spectra, b) cumulative eigenvalue diagram for the laminar jet ($Re \approx 709$, $We_l \approx 16$)

Vectors of the temporal changes of the first five POD modes are depicted in (Fig.2.12) where the indigo-colored line represents the constant stable nature of

the first mode whereas blue and black colored modes are showing the oscillatory behaviour of the next modes.

These vectors could be transformed into the Fourier space using the FFT command resulting in power spectral density (PSD) spectrum of the considered modes (Fig.2.13a). Critical frequency could also be extracted by analysing the energy of each of the DMD eigenvalues (Fig.2.13b).



(a) (b)

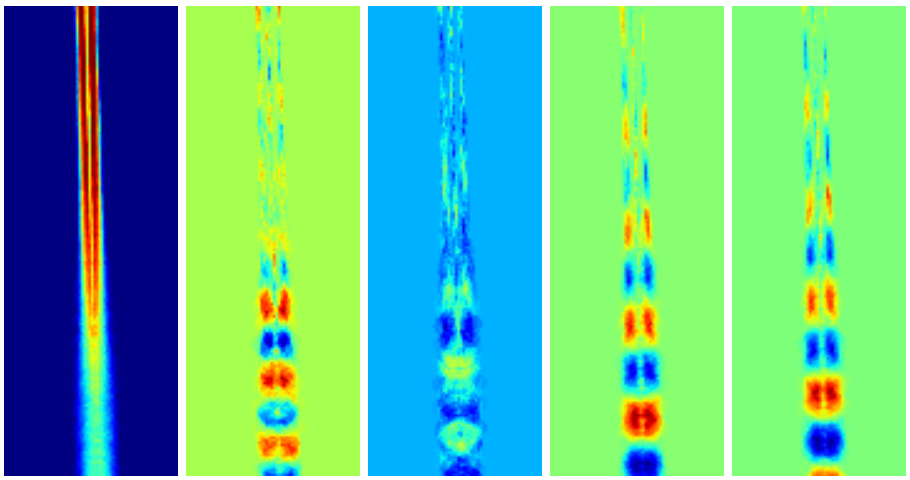
Figure 2.13: Modal frequency spectrum for the laminar jet ($Re \approx 709$, $We_l \approx 16$) a) POD-PSD, b) DMD energy

case 2: $Re \approx$ **1259,** $We_l \approx$ **51** [6]

By increasing the inlet jet velocity, we expect that the dominant spatial scales decrease alongside the increase in the amount of system frequencies. This can be exculsively performed by investigating the PSD for the most important POD modes (Fig.2.16a).

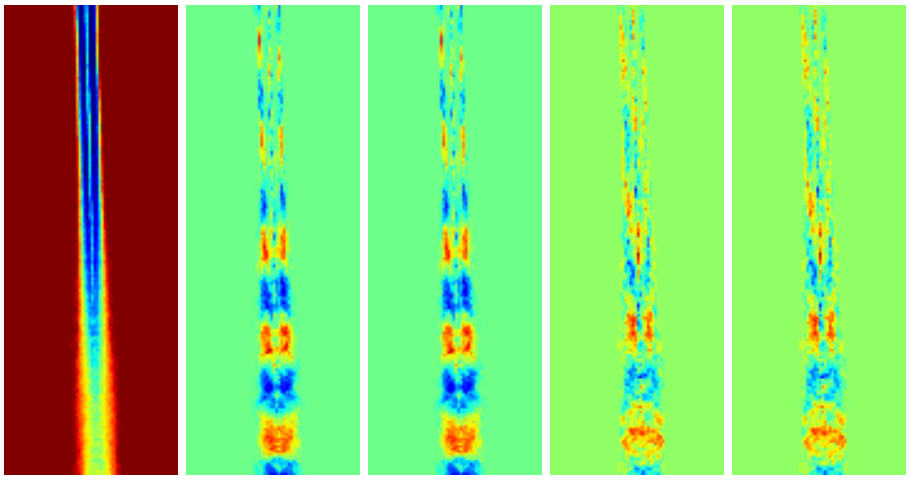
The first five POD modes are reshaped and illustrated in Fig.2.14.

[6] **The same setup as the one described in the section 4.2.1 is used here; liquid is water, orifice diameter D = 0.43 mm and the inlet velocity u = 2.93 m/s.**



(a) (b) (c) (d) (e)

Figure 2.14: First 5 POD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the laminar jet ($Re \approx 1259$, $We_l \approx 51$)



(a) (b) (c) (d) (e)

Figure 2.15: First 5 DMD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the laminar jet ($Re \approx 1259$, $We_l \approx 51$)

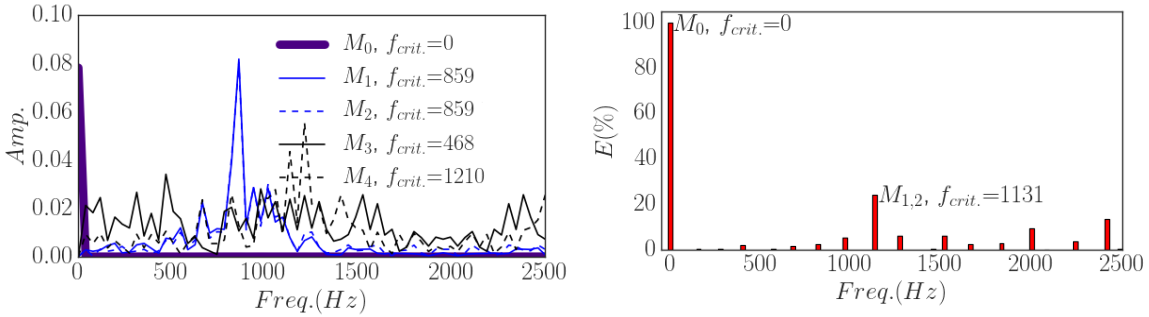
For this case, a delay in the occurrence of breakup could be observed as the breakup point height is lower compared to the previous case.

The selection of modes may be also discussed by considering both POD and DMD modal-decomposed patterns. Correlating the spatial structures of the modes

with the corresponding critical frequencies of the frequency domain, a better realization of the true modes is possible. A similar frequency content between the POD and DMD spectrum is visible where the fifth POD mode (Fig.2.16a) is well comparable with the the first and second DMD modes (Fig.2.16b). This is despite the fact that the first and second POD mode contain a higher energy amplitude. Here we may choose to put more weight on the fifth POD mode, however this is against the general idea of POD to arrange the modes based on an energy ranking scheme.

We also observe a critical frequency in the range of 2400∼2500 Hz in the DMD spectrum corresponding to the third and fourth DMD modes, that could be defined as a harmonic of the first and second modes (Leask and McDonell (2019)).

Generally, a shift in the values of the critical frequencies of the POD spectrum is observed with an increase of *Re* number. This is demonstrated by the fact that we extracted critical frequency content in the range of 507∼742 Hz for the first case but a range of 859∼1210 Hz for the second case. This range is both wider and larger in terms of frequency values for the second case (Fig.2.16a).



(a) (b)

Figure 2.16: Modal frequency spectrum for the laminar jet ($Re \approx 1259$, $We_l \approx 51$) a) POD-PSD, b) DMD energy

2.4.2 Liquid Jet In Crossflow

In this section, we analyze the behaviour of liquid jets in crossflow, which are the most practical systems in the industry. We categorize the jet in crossflow phenomenon into three domains with respect to the corresponding gaseous Weber

number defined as below:

$$We = \frac{\rho u_g^2 D}{\sigma} \tag{28}$$

where ρ is the fluid density, u_g is the gas velocity, d is the nozzle diameter and σ is the liquid's surface tension.

Momentum flux ratio could be also defined as the following:

$$q = \frac{\rho_l u_l^2}{\rho_g u_g^2} \tag{29}$$

The classification of jets in crossflow based on the corresponding Weber number was first done by Wu et al. (1997). We pick up a certain case for each of the flow regimes with specified flow conditions as below:

Table 2.2: Specification of the conditions for the selected datasets in the LJIC flow regimes

Case No.	Flow Conditions				
	q	We	u_l	u_g	Class
1	6.4	6.7	2.5	28.2	Capillary
2	11.3	20.5	5.8	49.3	Bag
3	17.2	50.6	11.3	77.6	Multimode

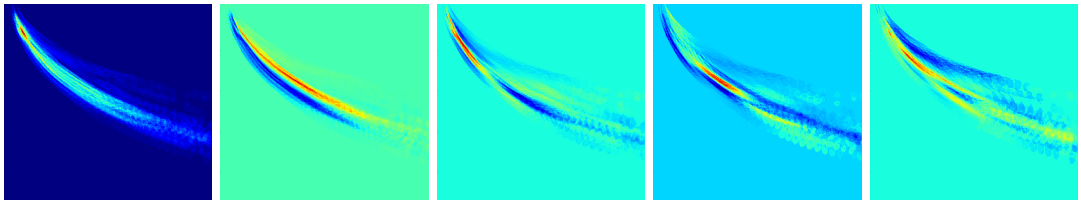
subscripts: $(l : liquid, g : gas)$
$[u]$ = m/s, D = 0.43 mm

(a) (b) (c)

Figure 2.17: Liquid jet in crossflow regimes: a) enhanced capillary, b) bag, c) multimode

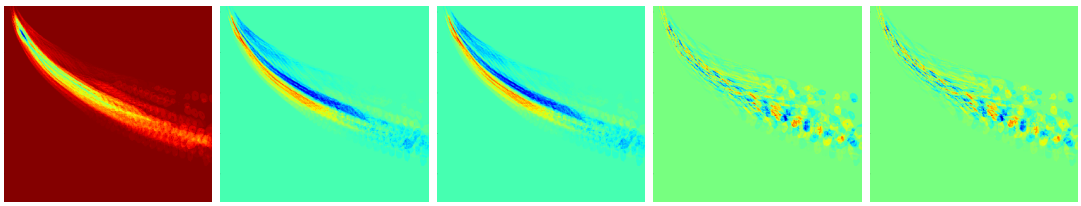
Enhanced Capillary Breakup

The laminar jet in crossflow is usually denoted as the enhanced capillary breakup regime, which is enforced by the gaseous crossflow, the jet bends and forms a curved trajectory. The first five modes for both POD and DMD schemes are shown in (Figs.2.18-2.19).



(a) (b) (c) (d) (e)

Figure 2.18: First 5 POD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the enhanced capillary breakup regime



(a) (b) (c) (d) (e)

Figure 2.19: First 5 DMD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the enhanced capillary breakup regime

Here we can observe that although POD method was successful in generating true modes for simpler cases like the laminar jet, applying it to the liquid jets in crossflow results in large spatial structures. Whereas with the DMD, we have extracted the fine scaled patterns in modes 3 and 4. A pile of regions with conjugate valued pixel intensities may be observed for both modes of 3 and 4 which is representative of the movement of the droplets after being detached from the liquid core.

Considering the POD and DMD modal energy spectra, a modal structure with frequency content in the range of 100∼200 Hz is detected for both methods. On the other hand, DMD extracts modes 3 and 4 with quite different structure than

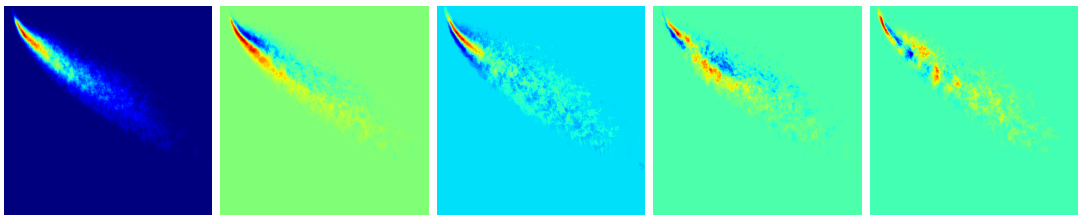
their POD counterparts. This is also explainable by comparing the PSD peaks with the DMD-driven frequency for the third and fourth mode that shows no common content.

Bag Breakup

By increasing the momentum flux ratio and the gaseous weber number, the enhanced capillary breakup regime shifts to the bag breakup regime mode. In this situation, some portions of the liquid get flattened and are forced by the cross-flowing air to detach from the core of the jet. The first five modes for both POD and DMD schemes are shown in Figs.2.20-2.21. First 5 POD modes are almost similar with the previous first five ones for the enhanced capillary regime although the occupied area by each of the modes is wider compared to its corresponding counterpart in the previous case. Despite having similar spatial structures, the difference between these two regimes could be identified through their PSD diagrams. As observed, a remarkable change in the values of the PSD peaks has happened and a wider area of frequencies is involved with the PSD diagram for the jet in the enhanced capillary breakup regime.

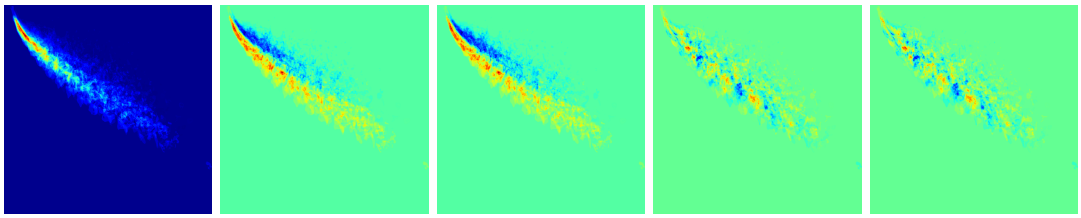
While the PSD spectrum offers that a more complicated system is under investigation for the bag breakup regime, this fact is not distinguishable from the modal spatial patterns. Manual selection of POD modes to better characterize the bag breakup regime is also not logical due to the essence of POD in automatically ordering the modes based on their energies which is quantified by their relative eigenvalues (this variable was defined in (9)).

On the other hand, DMD portrays structures of finer sizes which are characterized by their frequencies. For example, the second DMD mode identifies the similar structure as the second POD mode where a low-frequency displacement of the jet is captured. This displacement is in the form of a periodic oscillation where the jet is either at the maximum or at the minimum pixel intensity values at certain time-steps. This is better understood by calculating the time interval between the two consecutive snapshots where the jet is located at similar locations. This time interval is easily found by dividing the sampling frequency over the second mode's frequency ($f_s/f = 5000/756 \simeq 7$). It could be also verified by checking the location of the jet at every seven timesteps.



(a) (b) (c) (d) (e)

Figure 2.20: First 5 POD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the bag breakup regime



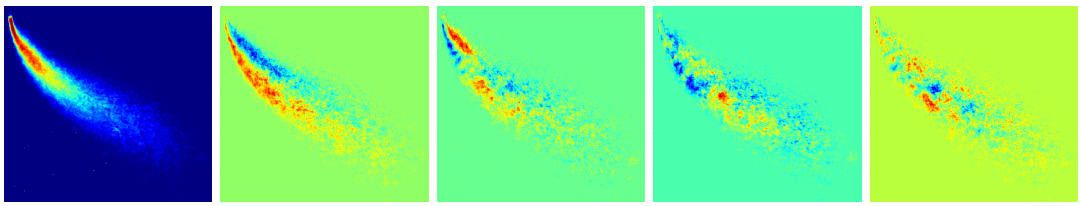
(a) (b) (c) (d) (e)

Figure 2.21: First 5 DMD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the bag breakup regime

Multimode Breakup

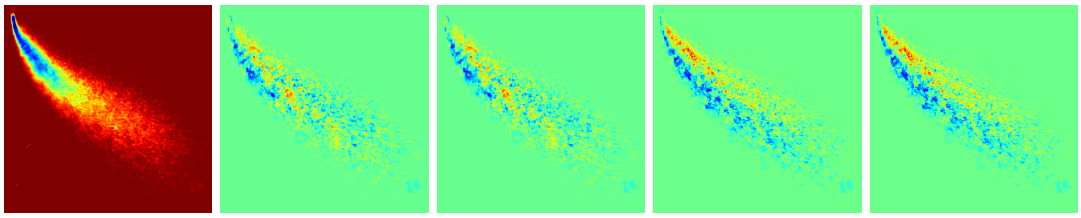
Further increasing the gaseous Weber number results in the shift of the breakup regime beyond the bag mode where the atomization is performed through both the bag breakup scheme and also the shearing of the small droplets off the jet core. The first five modes for both POD and DMD schemes are shown in Figs.2.22-2.23. The spatial patterns get even more distributed in terms of size with respect to the previous case. A more intricate system is expected here and is characterized using the POD and DMD modal frequency spectra (Fig.2.26).

First DMD modes shown in Fig.2.23 share significant content similarity with the first DMD modes of the bag breakup case. Except for the 0-th mode, all the presented modes here are identical in terms of spatial structures where it seems the third and fourth modes are the higher harmonics of the first and second modes.



(a) (b) (c) (d) (e)

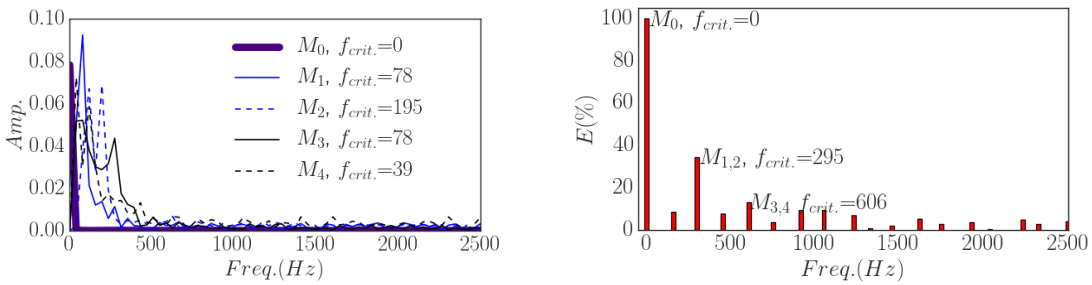
Figure 2.22: First 5 POD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the multimode breakup regime



(a) (b) (c) (d) (e)

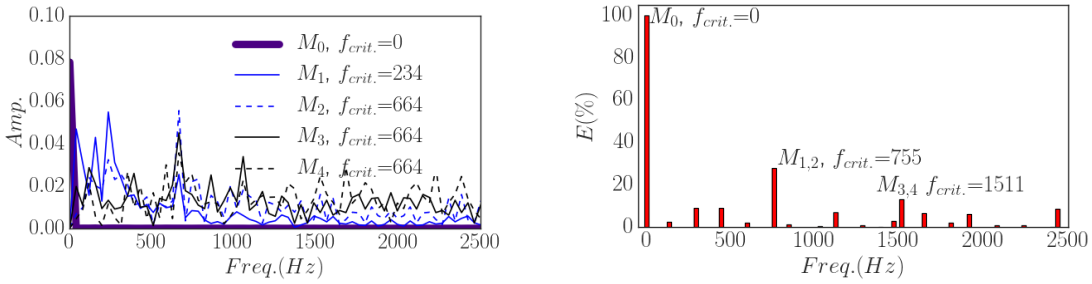
Figure 2.23: First 5 DMD modes, a) mode 0, b) mode 1, c) mode 2, d) mode 3, e) mode 4 for the jet in the multimode breakup regime

Frequency Spectrum



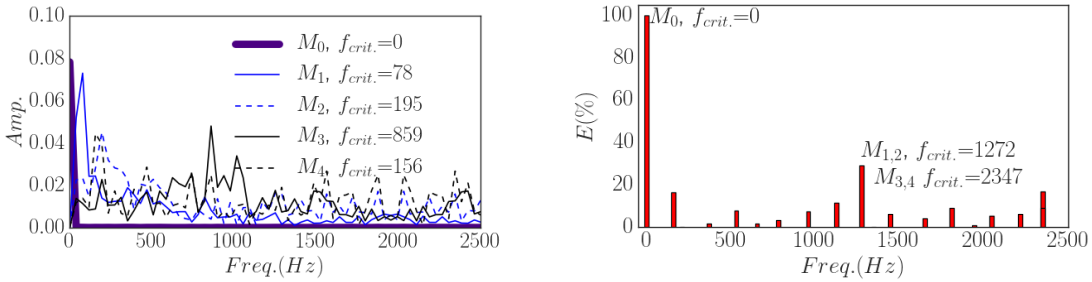
(a) (b)

Figure 2.24: Modal frequency spectrum for the enhanced capillary breakup regime a) POD-PSD, b) DMD energy



(a) (b)

Figure 2.25: Modal frequency spectrum for the bag breakup regime a) POD-PSD, b) DMD energy



(a) (b)

Figure 2.26: Modal frequency spectrum for the multimode breakup regime a) POD-PSD, b) DMD energy

In the table below, an overall estimation of the Strouhal numbers associated with the DMD modes for each of the cases studied above is presented. These values are all in the range of 0.06∼0.13 and correspond well with the values reported in the literature for the jet in the near-field (Eroglu and Breidenthal (2001)).

Table 2.3: Strouhal numbers associated with the near-field oscillations of the jet

Case	Dimensionless Numbers		
	Re	u_g/u_l	St
Enhanced Capillary	1252	11.3	0.06∼0.12
Bag	2900	8.5	0.07∼0.13
Multimode	5640	6.9	0.06∼0.10

There are a number of parameters which are highly dependent on the flow conditions like the gaseous Weber number and the momentum flux ratio. These parameters include the jet trajectory, penetration height and the column breakup point (CBP) location.

Jet Trajectory

Jet trajectory is usually defined as the path taken by the lower boundary of the liquid jet (windward side) (Wu et al. (1997)). While this characteristic parameter is usually found and depicted using image processing algorithms such as ImageJ and MATLAB curve-fitting tools, we are interested in characterizing this parameter using the modal spatial structures. For this purpose, POD or DMD first modes (average modes) could be utilized. These modes are representative of the average flow existent in all snapshots. A gradient function is applied on all pixels in order to find the liquid-gas interface and then the coordinates of the pixels with minimum -x- values in each row are reported as the windward trajectory of the jet.

A schematic of this procedure is depicted in Fig.2.27.

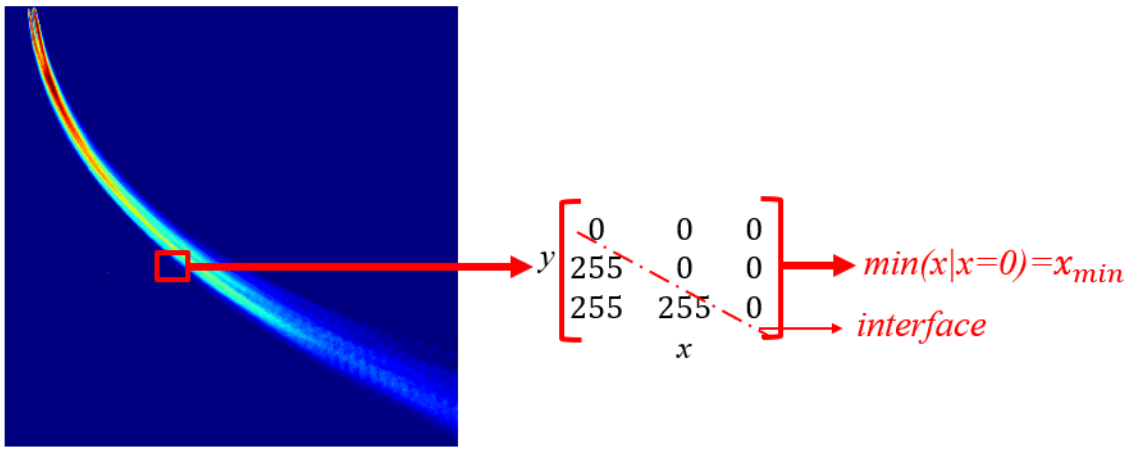


Figure 2.27: Average mode for the jet in the enhanced capillary breakup regime

Afterwards, a nonlinear least squares error curve fitting tool is used to predict a polynomial function in the form of the equation below for the trajectory:

$$\frac{y}{d} = a(q\frac{x}{d})^b \tag{30}$$

Constant values of a and b are found for each case, respectively. These values are comparable with the ones found by Wu et al. (1997).

Table 2.4: Jet trajectory characterization for different LJIC flow regimes

Case	Constants	
	a	b
Enhanced Capillary	1.39	0.49
Bag	1.38	0.50
Multimode	1.37	0.50
Wu et al. (1997)	1.37	0.50

2.5 Summary

In this chapter, an investigation to validate the efficacy of the POD/DMD methods has been conducted by generating a reduced order set in which each mode is representative of a meaningful spatial physics inside the phenomenon under study that also comes with its corresponding temporal content. Although the same approach seemed very unclear for describing the multiphase flow systems, it became a foundation upon which the extracted modes of laminar and crossflow jets may be interpreted. This explanatory study is the cornerstone for the next chapter where we examine the features of the elliptical jets and perform an analogical reasoning on their dynamic terms. The following chapter will also gain a great benefit from the discussions made in this chapter. A classification ML problem will be of interest where the features are the generated POD/DMD modes. It is highly important that a true set of modes get selected as the features because modes of higher harmonics do not add up to the available information and will be merely accumulated as an extra amount of data as previously discussed in this chapter.

Chapter 3

Dynamic Mode Decomposition of Elliptical Liquid Jets in Crossflow [1]

Abstract

Dynamics of round and elliptical liquid jets in subsonic crossflow is studied using high-speed imaging technique. The experiments are performed at constant gaseous weber number and liquid-gas momentum flux ratio of 6.45 and 17.87 respectively, with orifices of different aspect ratios having an equivalent diameter of 0.43 mm. All cases are carried out inside an open loop subsonic wind tunnel with a test section of 100×100×750 mm. For each case, dynamic modes are generated directly from the snapshots using a variant of Arnoldi method known as the dynamic mode decomposition (DMD). DMD results indicate that elliptical liquid jets have more small-scaled patterns with higher frequencies compared to the case of round liquid jets. As the first attempt to investigate the dynamics of elliptical liquid jets in crossflow, present work captures the dominant spatio-temporal structures. It is also found that the orifice aspect ratio can alter the jet wavelengths remarkably. The extracted data of this work can provide beneficial information on the behaviour of elliptical liquid jets exposed to the gas crossflow in the enhanced capillary breakup regime.

[1] **This chapter has been accepted to the Canadian Society for Mechanical Engineering International Congress (CSME 2020). Parts of it was submitted to the journal of Transactions of the Canadian Society for Mechanical Engineering (TCSME).**

3.1 Introduction

Liquid jet injection into a gaseous crossflow is an applicable mechanism in various fields such as thermal spray, ink-jet printing, and gas turbine engines (Wu et al. (1997)). Study of the near field behaviour of liquid jets in crossflow is essential since any oscillations on the jet body affects the functional efficiency of the spray. Identification of the principal phenomena could be useful in finding the coherent dynamic features and enhance the system performance as the liquid jet in crossflow (LJIC) is considered as a multi-scale phenomenon.

There are many governing parameters controlling the LJIC phenomenon. Wu et al. (1997) found the effect of liquid viscosity on liquid jet breakup. Taylor (1940) discovered the impact of gas density. Many other parameters were also investigated by analysing their effects on spray characteristics such as penetration height, column breakup point (CBP), Sauter mean diameter (SMD), spray spread angle in the works of Sallam et al. (2004); Lee et al. (2007); Herrmann (2010); Yunyi et al. (1998). Despite numerous researches in this area, only a few works have been performed on asymmetric nozzles due to their associated complexities. Among asymmetric nozzles, elliptical type orifices are investigated more than the others. Gutmark and Grinstein (1999) reported the enhanced mixing and spreading as the principal benefit of elliptical jets over their circular counterparts. One of the important features of elliptical liquid jets known as axis-switching is discussed in the work of Amini and Dolatabadi (2012). Study of such phenomena is often not straight-forward directly from raw experimental results such as time-resolved snapshots of high-speed camera. Therefore, flow structures may be analysed by decomposing them into modes using reduced order modelling (ROM) methods such as proper orthogonal decomposition (POD) (Weber (1931)), dynamic mode decomposition (DMD) (Holmes et al. (2012)), discrete Fourier transform (Rowley et al. (2009)), balancing modes for linear systems (Rowley (2005)), global eigenmodes for linearized dynamics (Bagheri et al. (2009)) and many other variants of these techniques.

Recently, Jadidi et al. (2019) showed that the flow structures at the exit of circular and elliptical nozzles are different and the elliptical nozzles show more prominent signs of turbulence. They also observed that for a specified momentum flux ratio and Weber number, the elliptical jets disintegrate earlier.

In this paper, ROM modelling of elliptical liquid jets exposed to low weber numbers in the enhanced capillary breakup regime close to the bag breakup regime is presented based on the DMD technique (Schmid (2010)). Using time-resolved snapshots, we extract the oscillatory physics of the flow and derive the relevant frequencies and the associated spatial modes using DMD. Dynamic mode decomposition (DMD) is an algorithm for finding Koopman modes and a standard variant of Arnoldi method (Rowley et al. (2009)). As a global stability analysis technique, DMD is applicable over a broad range of linear and nonlinear flows without any flow assumptions required (Mittal and Kumar (2003)).

Previous insight into the mechanism of LJIC was provided by an analysis technique based on proper orthogonal decomposition (POD) and spectral analysis (Arienti and Soteriou (2009)). DMD as the decomposition method used in this paper can be thought of as an ideal combination of spatial dimensionality-reduction techniques as POD with Fourier transforms in time. Each of the spatial modes in DMD are also associated with a given temporal frequency, possibly with a growth or decay rate (Kutz et al. (2016)). Thus, modal analysis of POD modes may result in association of each mode with different frequencies while DMD decomposes the input data into modes with singular frequency values for each of them.

3.2 Methodology

The experimental facility in this study consists of an open loop subsonic wind tunnel with a test section of $100\times100\times750$ mm, liquid injection system and an imaging system. A schematic of the experimental setup can be found in the work of Jadidi et al. (2019). A 1.5-HP blower fan made by Aeroflo is used to blow the ambient air into the wind tunnel. A Honeywell torque controller is utilized in order to adjust the air speed. The turbulence intensity of the gas velocity vectors was reported in the work of Farvardin (2013) where it was shown by particle image velocimetry (PIV) that the air velocity is constant and parallel in all regions except at the narrow boundary layers on the wall.

The liquid injection system also comprises of a pressure vessel, a flowmeter, and circular/elliptical injectors with different aspect ratios. Distilled water is the used liquid throughout the test. Injectors having plain orifices with length to

diameter of 10 are used where the equivalent diameter of different orifices is 0.43 mm. Aspect ratios are defined based upon the same rule provided in the work of Marzbali (2011) where the direction of the orifice major axis with respect to the crossflow direction (parallel/perpendicular) is considered as the criteria for the aspect ratio calculation. Elliptical orifices of aspect ratios 0.22 and 4.47, and a circular orifice of aspect ratio 1 are used in the present work. A tapered transition from 2 mm to 0.43 mm is considered in order to prevent cavitation effects inside the orifice (Farvardin (2013)). Mass flow rate of the liquid jet and the velocity of the wind tunnel of 38 g/min and 30 m/s are kept constant during all of the tests in order to ensure constant gas weber number (We) and momentum flux ratio (q) of 6.45 and 17.87, respectively.

The imaging facilities in this test include a Photron SA1.1 high-speed digital camera set at 5000 frames per second with a 1024×1024 resolution, a lens (AF Micro-Nikkor 105 mm f/2.8, Nikon) and two halogen lamps (type: T-3, 200T3Q/CL/78MM, Satco, USA).

The experimental tests were performed at atmospheric pressure and the room temperature and were repeated four times for each case at its unique aspect ratio. Fig.3.1 reveals three random snapshots of circular/elliptical liquid jets with different aspect ratios at the enhanced capillary breakup regime. For each of the cases, Dynamic Mode Decomposition (DMD) was applied on 500 consecutive snapshots for acquisition of coherent dynamical structures.

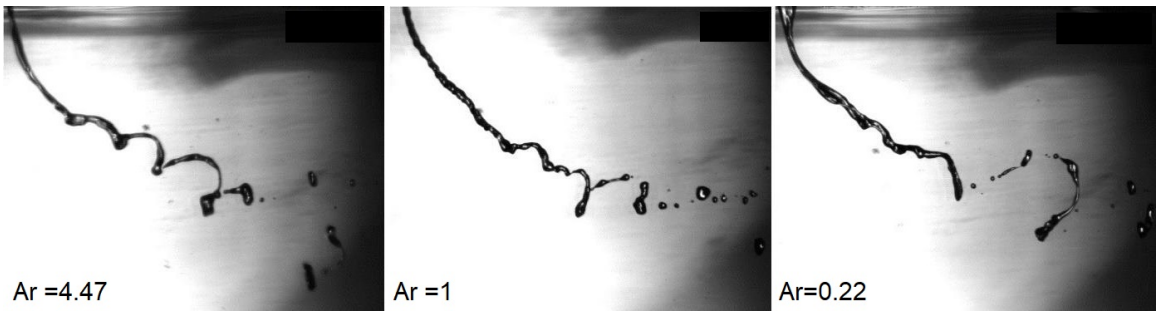


Figure 3.1: Enhanced capillary breakup of elliptical liquid jets; We=6.45, q=17.87

Using DMD, one can analyse the temporal coefficients of coherent spatial modes. Development of this algorithm to extract dynamic mode information from a flow field was carried out by Schmid (2010) based upon the Koopman analysis

of dynamical system (Rowley et al. (2009)). Current work is perhaps the first research on capturing the dominant dynamical modes to look at the small-scaled oscillations around elliptical liquid jets.

3.3 Results and Discussion

A low speed water jet exposed to a low speed gaseous crossflow in the enhanced capillary regime is investigated by DMD. As mentioned before, 500 snapshots of data were captured with a 5000 frames per second imaging speed for each of the circular/elliptical liquid jet cases among which 101 arbitrary snapshots were used in order to generate an A matrix with a maximum rank of 100 leading to the generation of maximum 100 dynamic modes. Furtherly, an r-truncation scheme (Kutz et al. (2016)) is developed in order to optimize the typical Ritz spectrum to have a near-equilibrium state of the system where the flow dynamics is almost linear. The exponential growth/decay rate and the stability/instability condition of the system could be discussed by the conjugate pair of eigenmodes. In addition, linear condition happens when almost all Ritz values lay on the unit circle $\| \lambda_i \| = 1$ that informs us about the coincidence of the sample points with an attracting set (Sarkar et al. (2013)). Similar observation of it was reported in the work of Rowley et al. (2009) for a liquid jet in crossflow (LJIC) case. As it can be seen in Fig.3.2, Ritz spectrum diagrams have been depicted for different values of r equal to 60, 80, 100 in case of having a circular liquid jet. there are different configurations of data points in order to track the effect of them on the stability of the reduced-order model.

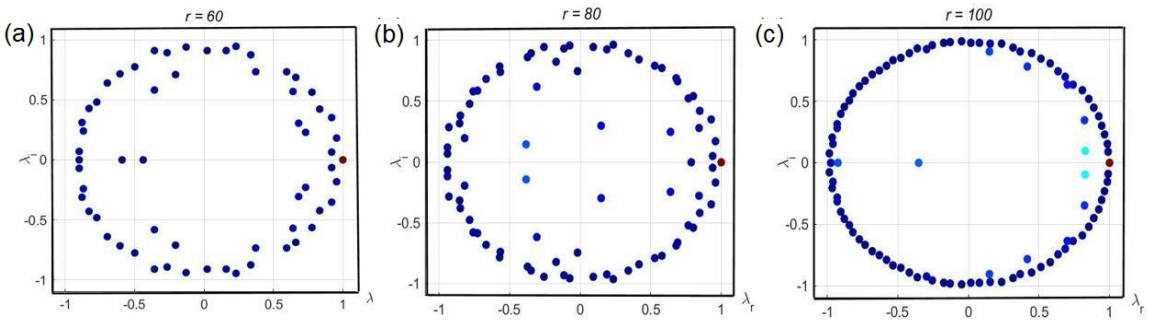


Figure 3.2: Ritz spectra for different r-truncation schemes; a) r=60, b) r=80, c) r=100

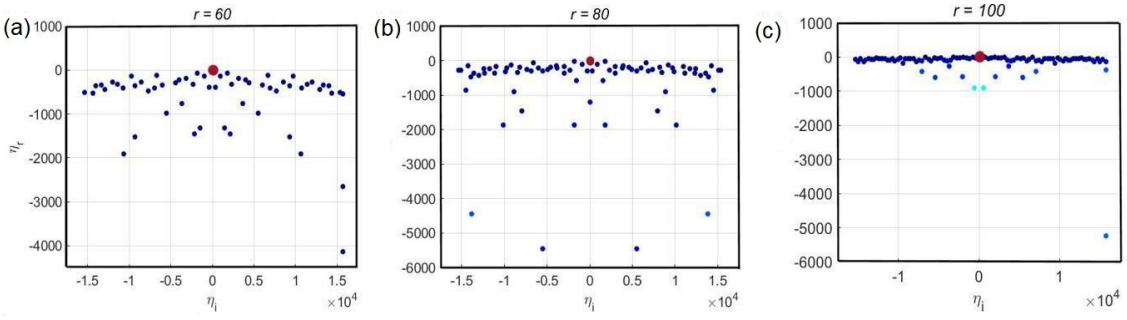


Figure 3.3: Global spectra for different r-truncation schemes; a) r=60, b) r=80, c) r=100

Hereby, an optimization approach based on trial and error was performed by maximizing the number of data points laying on the unit circle that guarantees the stability and also minimizing the energy content of the modes inside the unit circle. This approach has been applied on all circular/elliptical cases by trying different r values and an optimum r is found for each of the cases which is equal to 62, 72, and 67 for orifices with aspect ratios of 0.22, 1 and 4.47 respectively. The regarding Ritz and global spectra diagrams with the optimum aspect ratio for each of the circular/elliptical liquid jets is found in Figs.3.4-3.5.

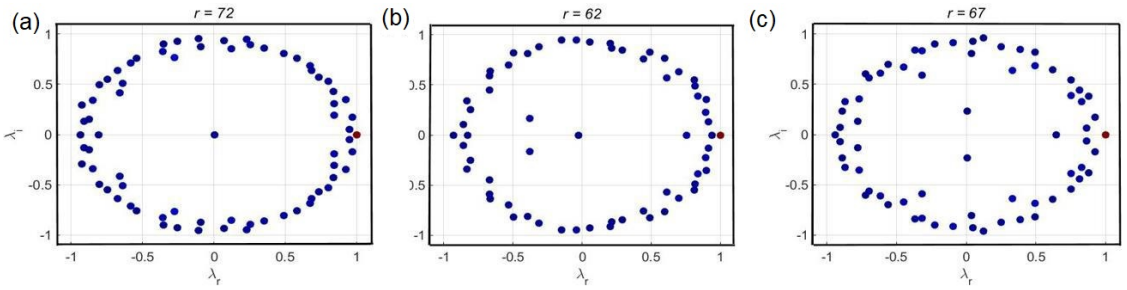


Figure 3.4: Ritz spectra; a) Ar=1, b) Ar=0.22, c) Ar=4.47 at optimum r values

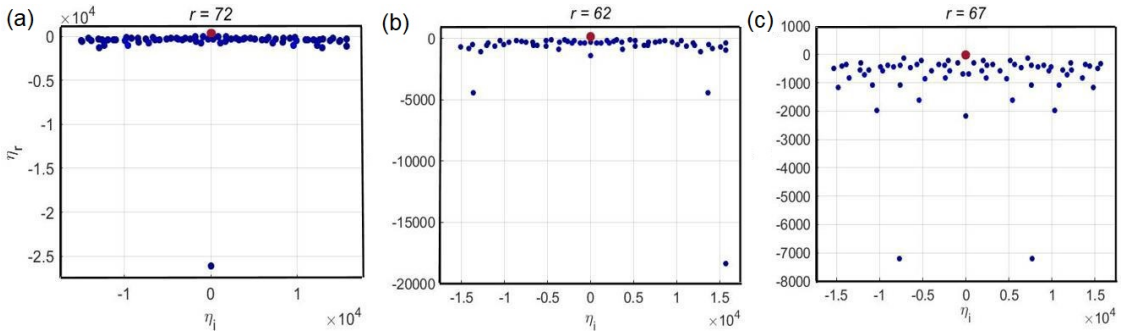


Figure 3.5: Global spectra; a) Ar=1, b) Ar=0.22, c) Ar=4.47 at optimum r values

In Figs.3.2-3.5, the filled circles in colours indicate the global energy norm of the associated modes. The red-coloured circle at $\lambda_i = 0$ and $\lambda_r = 1$ denotes the mean flow. It can be seen that for the case where r=100, two paired modes denoted by pale blue colour indicate that this model is not well suited since it is not linear and the decaying modes inside the circle have higher energy magnitudes with respect to their other counterparts laying on the unit circle. In general, the stability grows with increasing the value of r. However, higher values of r may result in modes with remarkable energy content laying inside the unit circle that is not desirable.

The principal modes of the circular/elliptical cases on the Ritz spectrum in Fig.3.4 are shown in a larger scale in Fig.3.6 where the associated modes are annotated with their indices.

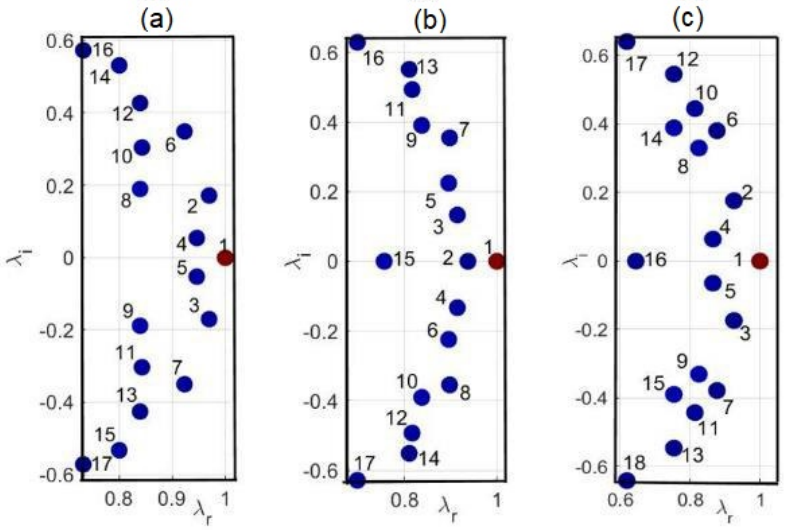


Figure 3.6: Ritz spectra for the mean and dominant dynamic modes at different aspect ratios; a) Ar=1, b) Ar=0.22, c) Ar=4.47

These modes dominate the entire energy spectrum with specified energy contents. Corresponding energy values are normalized with respect to the energy content of the mean mode (mode 1) using the definition of the form:

$$E = \frac{\| \phi_i \|}{\| \phi_1 \|} \tag{31}$$

The stability level of a single mode can be discussed based on its closeness to the unit circle as:

$$S = \sqrt{\lambda_{real}^2 + \lambda_{imag}^2} \tag{32}$$

From the Ritz spectra of dominant eigenvalues, the continuous oscillatory frequency can be found by the following relationship (Kutz et al. (2016); Sarkar et al. (2013)):

$$Frequency = \frac{imag(ln(\lambda))}{2\pi\Delta t} = \frac{arg(\lambda)}{2\pi\Delta t} \tag{33}$$

where Δt (the sampling time between two snapshots) and $arg(\lambda)$ are defined by the following expressions (Eqs.(34)-(35)):

$$\Delta t = \frac{1}{f_s} \tag{34}$$

$$arg(\lambda) = arctan(\frac{\lambda_{imag}}{\lambda_{real}}) \tag{35}$$

The sampling frequency is equal to 5000 in this case. Therefore, Δt will be 0.2 milliseconds. It is notable that the maximum capturable oscillatory frequency by DMD is limited by half of the sampling frequency.

Tables 3.1,3.2, and 3.3 show the real and complex eigenvalues of the mean and dominant modes, their associated absolute and normal energy content, stability S term for each of the modes, the angle of the modes with respect to the positive real axis and finally their corresponding frequencies. It is interesting that the dominant mode capturing the oscillatory behaviour of elliptical liquid jets in the enhanced capillary breakup regime at the threshold of bag breakup regime own higher frequency levels compared to the case of circular liquid jet. From Eq.(33), the average oscillatory frequency was calculated as 286 for the circular liquid jet whereas it was estimated as 349 and 343 for the major and minor axis elliptical liquid jets respectively (Table 3.4). These results confirm the existence of more small-scaled patterns with higher frequencies which is equivalent to smaller wavelengths associated with the elliptical jets with aspect ratios over or under 1.

Fig.3.5 portrays the global DMD spectrum for the circular/elliptical cases. The global spectrum for the circular liquid jet is presented in Fig.3.4a whereas Figs.(3.5b-3.4c) represent the same for the elliptical liquid jet. A strong alignment trend is capable to be considered along the line of the growth/decay rate $\eta_r = 0$. This trend is strongest in the circular case and the weakest in the elliptical case with the minor axis perpendicular to the direction of the crossflow. Strong alignment indicates a nearly linear model which is desirable. This issue could also be investigated through the S values where the closeness of each mode to the unitary circle is specified.

Table 3.1: Eigenvalues, energy content, stability and frequencies of the mean and dominant modes for the circular liquid jet (Ar = 1)

Mode No.	λ_{real}	λ_{imag}	$\parallel \phi \parallel$	E	S	$arg(\lambda)$	Frequency
1	1.00	0.00	125180	1.0000	1.00	0.00	0
2,3	0.96	$\pm$0.17	2221	0.0177	0.95	0.17	138
4,5	0.94	$\pm$0.05	2016	0.0161	0.94	0.05	44
6,7	0.026	$\pm$0.34	3285	0.0262	0.98	0.35	286
8,9	0.84	$\pm$0.18	5689	0.0454	0.86	0.22	176
10,11	0.84	$\pm$0.30	3207	0.0256	0.89	0.34	275
12,13	0.83	$\pm$0.42	1810	0.0145	0.94	0.47	374
14,15	0.80	$\pm$0.53	2313	0.0185	0.96	0.58	467
16,17	0.73	$\pm$0.57	868	0.0069	0.93	0.66	527

Table 3.2: Eigenvalues, energy content, stability and frequencies of the mean and dominant modes for the elliptical liquid jet with Ar = 0.22

Mode No.	λ_{real}	λ_{imag}	$\parallel \phi \parallel$	E	S	$arg(\lambda)$	Frequency
1	1.00	0.00	117920	1.0000	1.00	0.00	0
2	0.93	0.00	24	0.0002	0.93	0.00	0
3,4	0.91	$\pm$0.13	2861	0.0247	0.92	0.14	114
5,6	0.89	$\pm$0.22	3963	0.0336	0.92	0.24	196
7,8	0.89	$\pm$0.35	1900	0.0161	0.96	0.3765	299
9,10	0.83	$\pm$0.38	4410	0.0373	0.92	0.43	346
11,12	0.81	$\pm$0.49	2107	0.0179	0.95	0.54	432
13,14	0.80	$\pm$0.55	156	0.0013	0.97	0.59	475
15	0.75	0.00	5221	0.0443	0.75	0.00	0
16,17	0.69	$\pm$0.62	2295	0.0195	0.93	0.73	583

Table 3.3: Eigenvalues, energy content, stability and frequencies of the mean and dominant modes for the elliptical liquid jet with Ar = 4.47

Mode No.	λ_{real}	λ_{imag}	$\parallel \phi \parallel$	E	S	$arg(\lambda)$	Frequency
1	1.00	0.00	100070	1.0000	1.00	0.00	0
2,3	0.92	$\pm$0.17	2107	0.0211	0.94	0.18	148
4,5	0.86	$\pm$0.06	3483	0.0348	0.86	0.07	59
6,7	0.87	$\pm$0.37	1554	0.0155	0.95	0.40	323
8,9	0.82	$\pm$0.32	5135	0.0513	0.89	0.37	302
10,11	0.81	$\pm$0.44	2399	0.0240	0.92	0.49	397
12,13	0.75	$\pm$0.54	1048	0.0105	0.93	0.62	497
14,15	0.75	$\pm$0.38	5339	0.0534	0.84	0.47	378
16	0.64	0.00	2182	0.0218	0.64	0.00	0
17,18	0.61	$\pm$0.64	1770	0.0177	0.89	0.80	639

Table 3.4: Averaged stability and frequency of elliptical liquid jets based on their aspect ratios

Ar	$\bar{S}$	$\bar{Frequency}$
0.22	0.92	349
1	0.94	286
4.47	0.89	343

Based on Tables.(3.1-3.3), averaged S values of 0.94, 0.92, and 0.89 are estimated for the dominant modes in the circular/elliptical liquid jets the results of which can be found in Table.3.4.

In case of the energy norms, the values of E are given in Table.3.1 for the circular jet and in Tables.(3.2-3.3) for its elliptical counterparts. Modes with higher E play a more significant role in terms of contributing their share to the total energy of all modes in the reduced order model. For example, paired modes (8,9) in the elliptical case with Ar=4.47 own a 5.13% energy norm higher than any other paired modes of interest which make them the principal dynamics of the system. This result is validated by the depiction of the spatial structures of the real and imaginary parts of the principal modes in Figs.(3.7-3.8). These

modes were calculated and depicted for Ar=4.47 as the representative of all other studied cases. It is obvious from the real and imaginary parts of the modes that a 90° phase shift is existing between them. Strong similarity between the spatial structure of these modes with the original snapshot data depicted in Fig.3.1 explains their higher energy norm compared to the other modes.

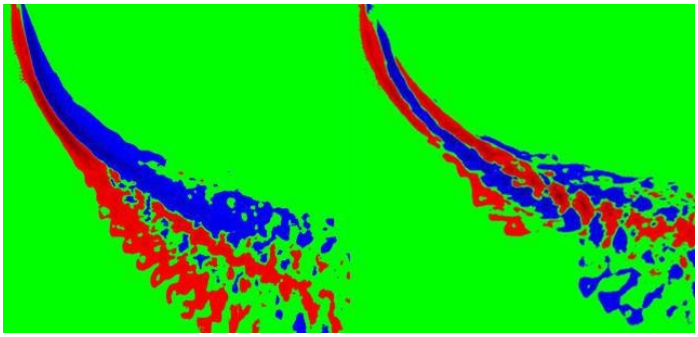


Figure 3.7: Real and complex parts of mode 2 for the elliptical jet with Ar=4.47

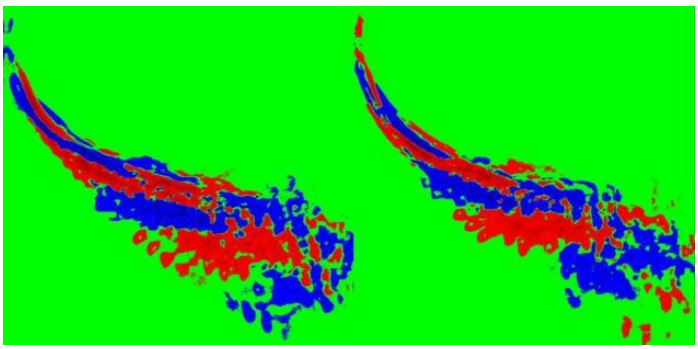


Figure 3.8: Real and complex parts of mode 8 for the elliptical jet with Ar=4.47

3.4 Conclusions

An investigation of the dynamics of the elliptical liquid jets in gaseous crossflows is performed at low weber numbers. Different elliptic-type orifices with similar equivalent diameter are used. Flow conditions are kept constant by enforcing the same air and liquid velocity for all tests. It is concluded that the flow structures coming out of the elliptical nozzles are totally different in comparison with their circular counterparts. For a given momentum flux ratio, the elliptical liquid jets contain more coherent small-scaled structures with higher frequencies compared to the circular jets. Higher frequencies lead to lower wavelengths that means faster disintegration and less penetration into the crossflows for the elliptical jets which is perhaps due to the effects of drag force, axis-switching and the turbulence. This conclusion is highly consistent with the previous findings in the work of Jadidi et al. (2019).

Chapter 4

Supervised Classification of Liquid Jets in Crossflow [1]

Abstract

Breakup of liquid jets in crossflow (LJIC) contain unique embedded patterns based on the type of the pertained flow regime. Recognition of these patterns and correlating them to the underlined flow schemes is a possible but yet challenging task due to their complex nature. In this research, we have utilized unsupervised reduced-order models to create a feature-based supervised classifier that diagnoses multiple flow regimes. These models include proper orthogonal decomposition (POD), principal component analysis (PCA) and dynamic mode decomposition (DMD). Snapshots are being extracted by high-speed imaging of the flow field of fourteen different cases at various categories. These images are then stacked into a big-data matrix as the train set for the support vector machine (SVM) and random forest (RF) classifiers to learn. Then, the generated classifiers in the previous step are used to predict which category belongs to every dataset of the six newly imported cases. Afterwards, the accuracy level of different permutations of reduced-order models and machine learning algorithms is calculated. Results indicate that using dynamic modes of DMD in partnership with the RF algorithm outperforms every other model with the highest accuracy rate of 95%. Finally, a decision-maker application that classifies the datasets based on the first

[1] **This chapter has been submitted to the journal of atomization and sprays and is currently under review.**

three models with the highest accuracy levels is introduced to provide a user-friendly environment for data classification at all other potential conditions.

4.1 Introduction

The breakup processes of liquid jets in crossflow include enhanced capillary, bag, multimode and shearing regimes. These processes commence with the deformation and flattening of the liquid column, and the growth of surface waves. Consequently, the liquid column is disintegrated into ligaments and droplets. When the gaseous Weber number (We) is below 10, the liquid jet is curved by the aerodynamic forces that reinforce the capillary forces while being under the control of the surface tension forces. The corresponding breakup regime to this behaviour is called the enhanced capillary breakup (Kitamura and Takahashi (1976)). As the gaseous Weber number increases, the liquid column undergoes breakup behaviours similar to the secondary breakup of spherical droplets (Krzeczkowski (1980)). These behaviours lead to the formation of the bag, multimode and shear breakup regimes based on the magnitude of We number.

The breakup regimes in the liquid jet in crossflow (LJIC) configurations do play a significant role in determining the physics of the jet at both the primary and secondary scales (Wu et al. (1997); Mazallon et al. (1999); Sallam et al. (2004); Lee et al. (2007)). These processes affect the droplet size and velocity distributions, penetration depth, and the breakup point location each of which is a game-changer in several sectors of industry such as delivering fuel in engines, burners, boilers, agricultural sprays, etc. (Jadidi et al. (2016); Curran (2001)).

Assuming LJIC phenomena as a nonlinear dynamical system, analysis of such processes seems crucial by the extraction of spatiotemporal patterns from the data generated by numerical simulations or experiments. To reach this goal, several reduced-order modelling techniques, all of which are founded based on the singular value decomposition (SVD) of data, have been introduced in the fluid mechanics, mathematics and computer science communities. It could be known as a generalization of the Fourier transform (Brunton et al. (2020)). SVD has many applications mainly focused on data reduction, dimensionality reduction and has been promoted as a foundation for machine learning in recent years. The most

well-known models in this area are principal component analysis (PCA) (Pearson (1901)), proper orthogonal decomposition (POD) initially introduced to the fluidics community (Berkooz et al. (1993)), and dynamic mode decomposition (Schmid (2011)).

POD and PCA are both static models analogous to each other in the sense that they have similar definitions except that PCA uses the mean-subtracted data as the input. On the other hand, DMD is a dynamic dimensionality reduction model with the capability to extract dynamic information from flow fields. In both frameworks (static vs. dynamic), the output data is usually in the form of modes, each of which is representative of a specific flow regime. The physical mechanism embedded in the modes is on a dynamical system of importantly, fewer degrees of freedom compared to the initially given system (Schmid (2011)). The distinction between these two frameworks relies on the fact that POD and PCA determine the optimal set of modes to represent the data based on the energy norm. At the same time, DMD captures dynamic modes with associated growth rates and frequencies.

As mentioned previously, POD was presented to the fluidics community in 1993 by Berkooz et al. (1993). They described the method, illustrated its use in the analysis and modelling of turbulent flows. Bernero et al. studied the jet in counterflow phenomena by investigating the generated modes using POD (Bernero and Fiedler (2000)). Meyer et al. analyzed the experimental data of a turbulent jet in crossflow using POD and found that the shear layer vortices are not coupled to the dynamics of the wake vortices (Meyer et al. (2007)). Also, hanging vortices were identified, and their contribution to the counter-rotating vortex pair (CVP) was described. Arienti et al. linked the observed travelling waves in the LJIC systems to the fastest growing wave of Kelvin-Helmholtz instability using the frequency and wavelengths obtained by applying POD (Arienti and Soteriou (2009)).

In 2009, Rowley et al. (2010) addressed all previous developments in model-reduction techniques applicable to fluid flows and introduced a new method for analyzing nonlinear flows based on spectral analysis of the Koopman operator, a linear operator defined for any nonlinear dynamical system. They showed that, for an example of a jet in crossflow, the resulting Koopman modes decouple the

dynamics at different timescales more effectively than POD modes and capture the relevant frequencies more accurately than linear stability analysis. This work became a foundation upon which Schmid established dynamic mode decomposition (DMD) and used it to break a fluid process into dynamically relevant and coherent structures that aid the characterization and quantification of physical mechanisms in the fluid flows (Schmid (2011)). In 2014, Tu and Rowley (2014) performed an analysis of 3D separated turbulent flow over a finite-thickness plate with an elliptical leading edge at Re=100,000. They compared the final POD and DMD modes in terms of their ability in data reconstruction and frequency capturing. In addition to the fluidics community, DMD has gained a great deal of attention in the market strategies (Mann and Kutz (2016)), neuroscience (Brunton et al. (2016)), disease recognition (Xi and Zhao (2019)), etc.

In the current work, we intend to choose different reduced-order modelling techniques as the feature-extraction tools and furtherly build a machine-learnt classifier that diagnoses different flow breakup regimes (enhanced capillary, bag, and multimode) using image snapshots. Previous studies in this area are limited to the classification of data-driven thermal fluid models (Chang and Dinh (2019)), classification of boiling regimes (Hobold and Da Silva (2018)), characterization of two-phase flows (Chakraborty and Das (2020)), and flow pattern classification in liquid-gas flows using flow-induced vibration (Carvalho et al. (2020)). The objective of this paper is to compare the accuracy of different models in predicting the breakup regime of a new LJIC case based on what is learned in the training step using the image datasets of various LJIC cases with their known breakup categories. Two classic machine learning algorithms of SVM and RF are used in this work to generate the desired classifiers. SVM constructs a set of hyperplanes in a high dimensional space, which can be used for classification, regression, etc. RF is an ensemble learning method for classification, regression, and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Using other machine learning and deep learning approaches as convolutional neural networks (CNNs) that are powerful in image classification was also considered but found to be unsuitable due to the insufficient number of images (Rawat and Wang (2017)).

The remaining of this research paper is organized as follows: In the second part, the experimental setup and the data collection methods are explained. Then, discussing the methodology and the governing equations behind each model is included in the third part. Finally, the results are generated and addressed by implementing the methods in the fourth part.

4.2 Methodology

4.2.1 Experimental Setup

The experimental facility in this study consists of a closed-loop subsonic wind tunnel with a test section of (200$\times$ 200$\times$ 400 mm), a liquid injection setup, and a high-speed image capturing system. The test section is made of plexiglass material for better visualization. A schematic of the setup is shown in Fig.4.1. A blower fan made by ABB is used to blow the air into the wind tunnel. The air velocity distribution, as well as its turbulence intensity, were measured by particle image velocimetry (PIV), and a uniform profile was obtained in all regions except at the 5-mm thin boundary layer on the wall at the injector location (Farzad (2019)). The turbulence intensity was also calculated to be 5% at the same location. The liquid injection system consists of a pressure vessel, a flowmeter, and a single injector with an orifice diameter of 0.5-mm. It is also notable that a tapered transition from 2 mm to 0.5 mm is considered to prevent cavitation inside the orifice. With DI water as the test liquid, the air crossflow and injection velocities are varied over a wide range to provide an extensive dataset of experimental results. The liquid velocity ranges between 2.5 to 11 m/s, whereas the considered velocity range for the gaseous crossflow is 28 to 84 m/s.

The image capturing system is based on the backlighting technique. It includes a Photron SA1.1 high-speed digital camera set at 5000 frames per second with a 1024$\times$1024 resolution for all cases, a lens (AF Micro-Nikkor 105 mm f/2.8, Nikon), and two halogen lamps (type: T-3, 200T3Q/CL/78MM, Satco, USA).

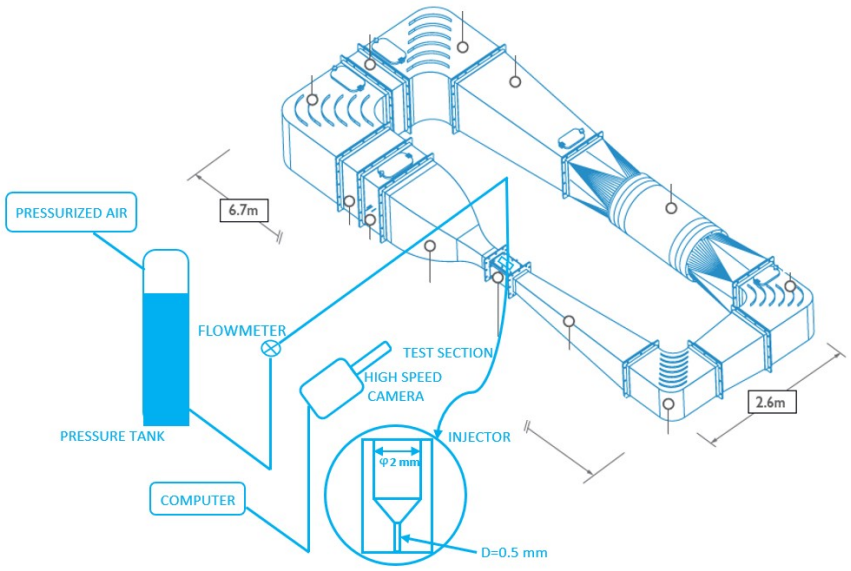


Figure 4.1: Schematic of the experimental setup

4.3 Data Collection

In this study, we used twenty arbitrary LJIC cases with a test size equal to 0.3 (fourteen cases for the training phase and six cases for the test phase), all located in the triple-typed flow regimes of the enhanced capillary, bag and multimode breakup, the details of which are shown in Fig.4.2. It is highly noteworthy that the existence of surface/column breakup has been neglected here as the amounts of momentum flux ratio are adequately small. The original map could be found in the work of Wu et al. (1997). The corresponding amounts of momentum flux ratio and the gaseous weber number are listed in Table 4.1.

Table 4.1: Specification of the LJIC cases

Case No.	1	2	3	4	5	6	7	8	9	10
q	6.4	34.5	130.1	4.1	22.1	83.5	2.8	15.3	58.2	2.1
We	6.7	6.7	6.7	10.5	10.5	10.5	15.1	15.1	15.1	20.5
Case No.	11	12	13	14	15	16	17	18	19	20
q	11.3	42.6	8.6	32.6	6.8	25.8	4.5	17.2	3.9	14.9
We	20.5	20.5	26.8	26.8	33.9	33.9	50.6	50.6	58.4	58.4

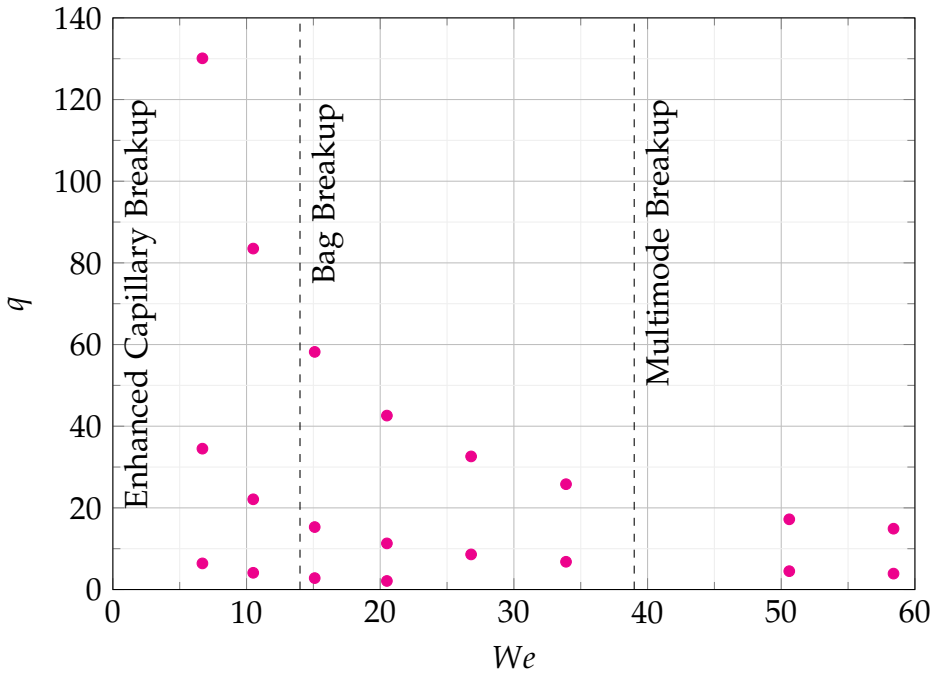


Figure 4.2: Visualization map of the LJIC cases

As indicated above, a database consisting of 500 consecutive images at twenty different cases (25 images per case) has been chosen to generate the necessary patterns. Each picture is converted to a tall skinny columnar matrix with a hexadecimal number system (pixels with values 0∼255) and attached to its peer matrix of the next image. This process is performed on both the train and the test sets. We consider distinct LJIC cases located in different flow regimes, the details of which are shown in Fig.4.2. A better understanding of the whole process is shown in Fig.4.3.

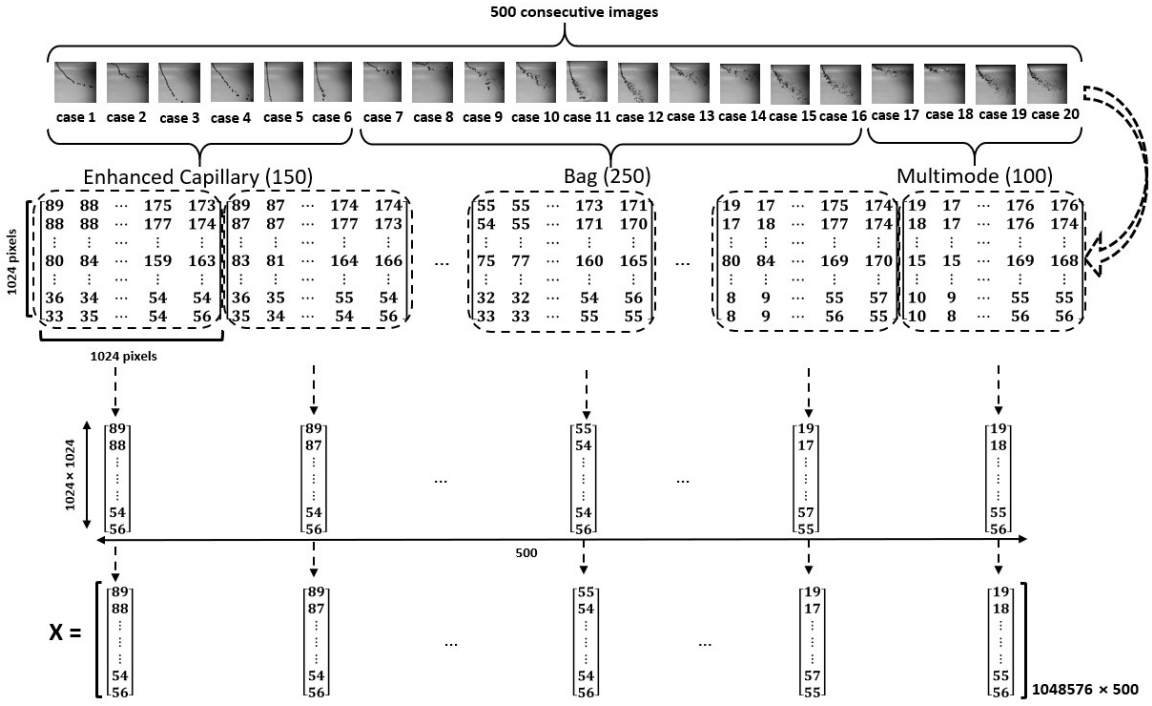


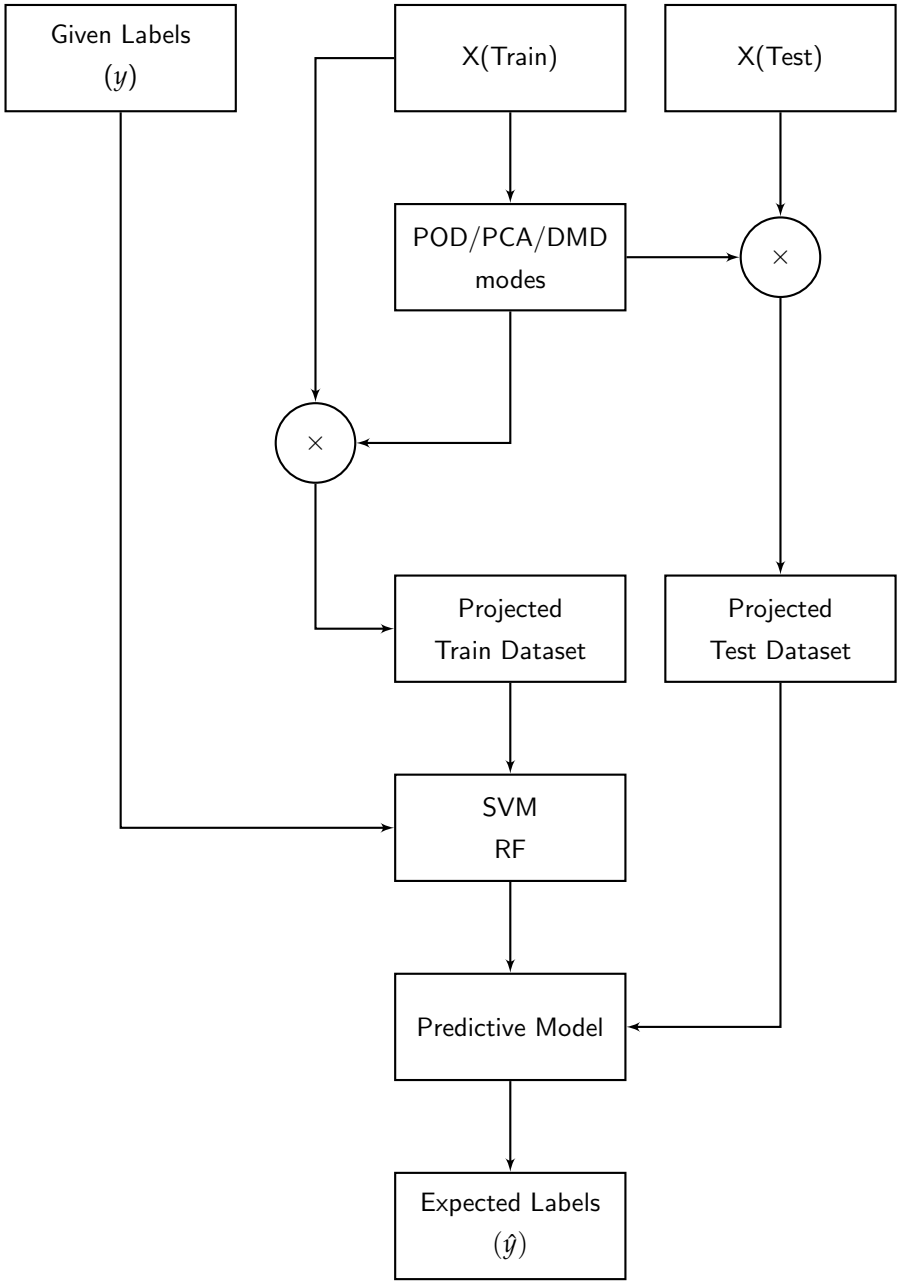
Figure 4.3: Data collection procedure

4.3.1 Supervised Learning Pipeline

The full data matrix X is now ready to be imported into various SVD-based algorithms ie. POD, PCA, and DMD. As indicated previously, a test set size of 0.3 is considered for all iterations. As shown in Fig.4.4, the train set is first used to build the features. Then, both train and test set images are projected into the features to generate the projected data matrix that has a (500$\times$rank) dimension. The truncation rank can be modified manually or be set to the optimal hard threshold for singular values found by Gavish and Donoho (2014). This method is used to recover the low-rank matrices from noisy data by hard thresholding the singular values obtained from reduced-order models such as POD, PCA and DMD. For the sake of simplicity, auto-rank selection based on this method is used here and is found to be equal to 12 for the current train dataset consisting of 350 consecutive images at 14 different cases. It means twelve modes are kept in each reduced-order model. Afterwards, both train and test image datasets (collected from images) are projected into the available previously generated modes that have the same dimension as the images (1024$\times$1024). The projected output data

for the train and test sets will be a dataset with a dimension equal to the number of input images in each of those datasets times the number of retained modes. This amount is equal to (350×12) here for the train and (150×12) for the test set, respectively. In the next step, the projected data obtained from the train set is imported along with the corresponding labels (A vector with a size equal to the number of input images containing the category type info of each image) into the machine learning model (SVM or RF), and the predictive model is generated in the output. In the last stage, the projected data obtained from the test set is given to the predictive model, and a vector of expected labels specifying the category type of each image is created. The entire procedure is performed for the permutation of three reduced-order models i.e. POD, PCA, and DMD with the two machine learning methods; SVM, and RF, thus making six different labelling vectors. Each of the six predicted labels is obtained by repeating the whole pipeline for 10 times to get better classification rate results. A ten-fold cross-validation is also used for error analysis as described in the work of Xi and Zhao (2019).

Figure 4.4: Flowchart of the machine learning pipeline



In the next section, we review the mathematical equations of POD, PCA and DMD as well as discussing some fundamentals of SVM and RF.

4.3.2 Methodology and governing equations

Reduced-Order Modelling

This section can be separated into two subsections: static and dynamical modelling. Static models like POD and PCA both use SVD of data to generate the feature modes. The only difference between POD and PCA lies in the fact that PCA uses the mean-subtracted data as the input. In contrast, dynamical methods such as DMD perform this task with three additional steps compared to the static ones. DMD first computes the correlation matrix that maps the first (N-1) images to the second (N-1) images in the dataset using the static modes obtained by performing SVD on the first (N-1) images. Then the dynamic modes are found by solving the eigendecomposition problem for the correlation matrix. A deeper insight into the details of this procedure could be found in the recent work of Brunton and Kutz (2019). It is notable that the columns of the input matrix X are proceeding in time (Fig.4.5) within one case but have abrupt change from one case to another. These cases are correlated by the Weber number and momentum flux ratio, although they are not connected temporally.

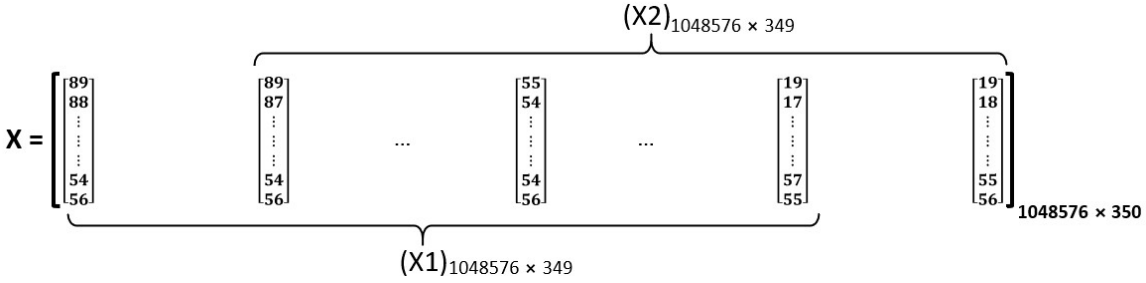


Figure 4.5: Schematic of the train set columns' time proceeding in DMD

A brief mathematical review of different reduced-order modelling techniques is given in (36)-(37)-(38):

$$X = U\Sigma V^* \Longrightarrow U : POD\ modes \tag{36}$$

$$\hat{X} = X - \bar{X} \Longrightarrow \hat{X} = U\Sigma V^* \Longrightarrow U : PCA\ modes \tag{37}$$

$$(X1) = U\Sigma V^* \Longrightarrow \tilde{A} = U^*(X2)V\Sigma^{-1} \Longrightarrow$$
$$\tilde{A}W = W\Lambda \Longrightarrow \varphi = (X2)V\Sigma^{-1}W \Longrightarrow \varphi : DMD\ modes \tag{38}$$

In (36), POD takes advantage of the singular value decomposition technique to decompose the X matrix into its constituent parts. U denotes the matrix of spatial embedded patterns, Σ remarks the eigenvalues containing the energy of each pattern and the V matrix specifies the temporal changes of the modes. In (37), PCA has a similar approach with the same notation for the eigenmodes, eigenvalues and the temporal terms. In (38), DMD solves the eigenvalue problem of the correlation matrix $\tilde{A}$ and reconstructs the spatial modes φ based on that. A full understanding of the methods is provided by Brunton and Kutz (2019).

Machine Learning Models

Two supervised classical learning methods of SVM and RF are used in this study due to their superb functionality in classifying datasets (Noble (2006); Liaw et al. (2002)). A Gaussian kernel has been used for the SVM, and a ten-fold cross-validation approach for both models is remarkable to minimize the probable over-fitting of the model. One thousand decision trees are used for the random forest classifier. The scikit-learn Python library is utilized for training and testing the classifiers and building the predictive model. The overall accuracy for each model is then calculated as below (39):

$$Accuracy = 1 - \frac{missclassified\ samples}{total\ samples} \tag{39}$$

The reduced-order modelling and machine learning parts are all designed and programmed in object-oriented Python language. We also used the cross-platform GUI toolkit named Qt to provide the graphical user interface application version for better user-friendliness. This application is easily accessible through a Github repository called PySVD [2].

[2]The data and code corresponding with this work is available online at: https://github.com/k1mokhtarpour/pysvd

4.4 Results and discussion

4.4.1 Feature Extraction

In this section, we provide a trivial introduction to the interpretation of features generated by static and dynamic reduced-order modelling techniques. The liquid jets are broken up and stretched through the interaction with the cross coming airflows. The corresponding embedded patterns are correlated with the type of breakup regime. All singular values are ordered in a descending logarithmic amplitude scheme where the first twelve modes are truncated using the optimal hard threshold method for singular values (Gavish and Donoho (2014)). The logarithmic singular values distribution versus the number of modes is displayed in Fig 4.6 for the first 250 modes, where a steep reduction is seen for the first few values before reaching a near plateau state. This figure shows that nearly all of the energy relies on the first twelve modes. Therefore, twelve modes are retained for each of the reduced-order models of POD, PCA and DMD.

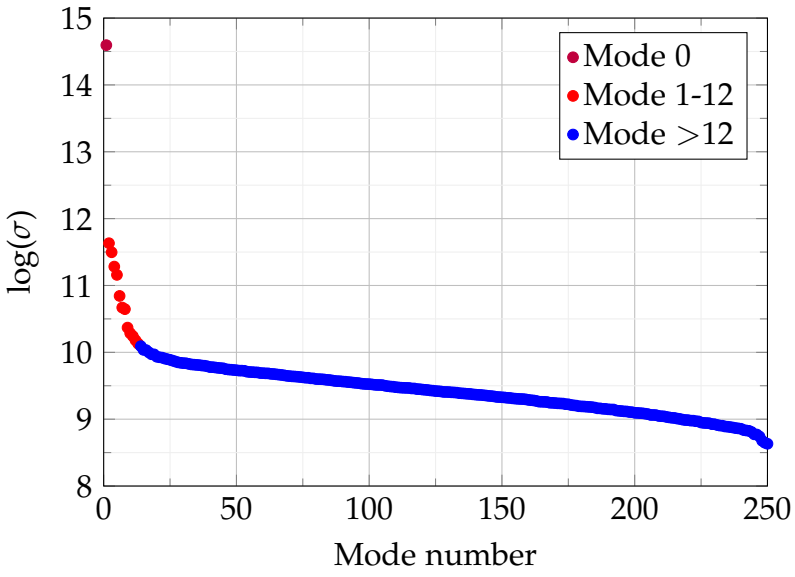
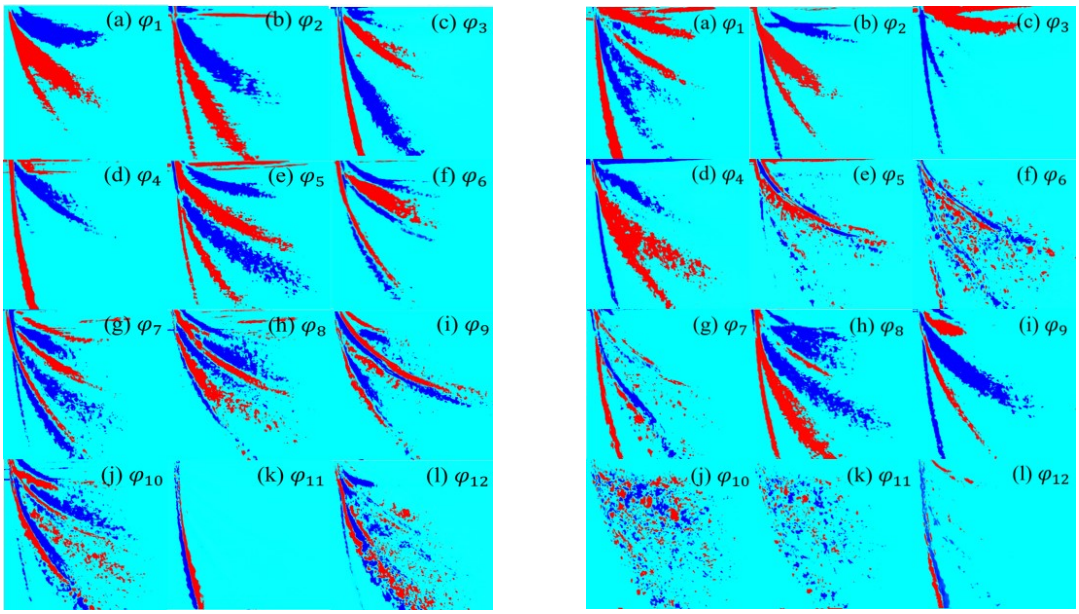


Figure 4.6: Logarithmic singular value distribution versus the mode number

The first twelve modes in both static (POD) and dynamic (DMD) models are depicted in Fig.4.7, where the average mode called mode zero, is removed in

both cases. All depicted features are a combination of different physics occurring at different flow conditions since we considered a big data matrix containing the information of LJIC systems at fourteen various conditions. Therefore, it is quite hard to assign any of the features to a single type of flow. Despite this, we can still observe some modes that are closer to the behavior of one regime. As an example, here, in the case of the enhanced capillary breakup regime, the features are cleaner with fewer oscillations as the breakup occurs later, and the jet penetrates more compared to bag and multimode schemes with the same gaseous weber number.



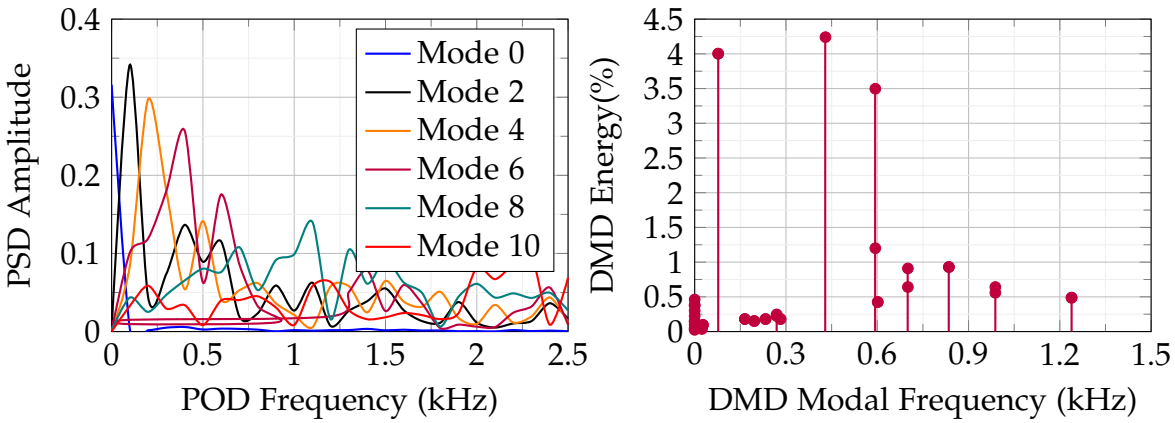
(a) (b)

Figure 4.7: First 12 dominant modes, a)POD, and b)DMD

For example, POD mode 11 shows the jet penetration fluctuation of the jet where the red and blue pixels are representing the maximum, and minimum values versus the zero-valued paled blue pixels. This result agrees well with the most jet trajectories in the enhanced capillary case number 3. The similar mode in DMD (mode 12) is contaminated with some small patterns at different locations. This observation shows DMD localizes less dominance to the features of single flow regimes compared to POD while trying to decompose the dataset in terms of the most spectrally essential patterns. It means DMD adds up elements of the

same frequency from different flow categories, while POD generates an energy-based ranking of modes. That is why more trajectories could be seen in the first DMD modes in comparison with the first POD modes.

Dynamic models could be compared with the static models in the frequency domain. The power spectral density (PSD) diagram of some POD modes, as well as the energy versus the modal frequency of DMD modes, is observable in Fig.4.8. It clearly shows that the PSD amplitudes extracted from the temporal term (V) vectors contain peaks at various locations. It means that each POD mode is corresponding with a few numbers of frequencies, while each DMD mode includes a single frequency. This fact shows that DMD outperforms POD in the sense that it distinguishes the exact frequencies, while only a spectrum of frequency could be found using the fast Fourier transform (FFT).



(a) (b)

Figure 4.8: Frequency spectrum, a)POD, and b)DMD

4.4.2 Supervised Classification

In this section, we consider different combinations of reduced-order models with classification algorithms. As mentioned earlier, dynamic features are expected to predict a higher number of samples accurately compared to the static features. To investigate this issue, the accuracy levels of different models is depicted in Fig.4.9, where the box plots show the accuracy variance for each of them.

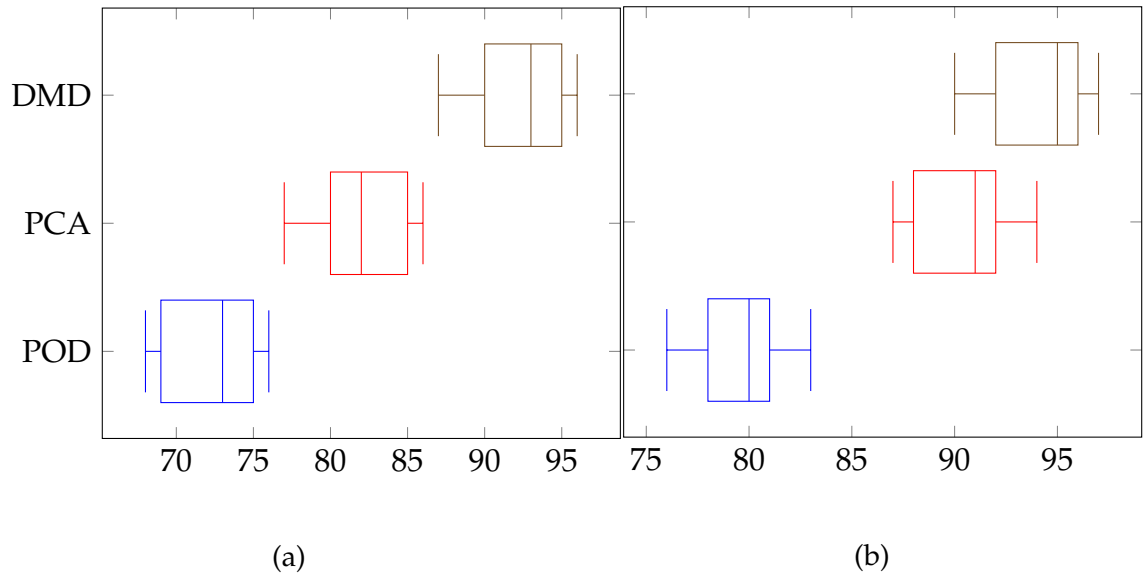
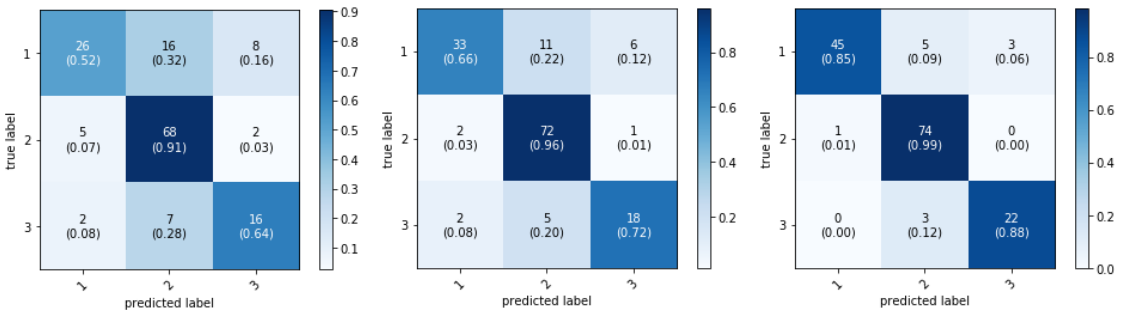


Figure 4.9: Box plot of classification rates using different combined reduced-order modelling and machine learning algorithms; (a) SVM and (b) RF

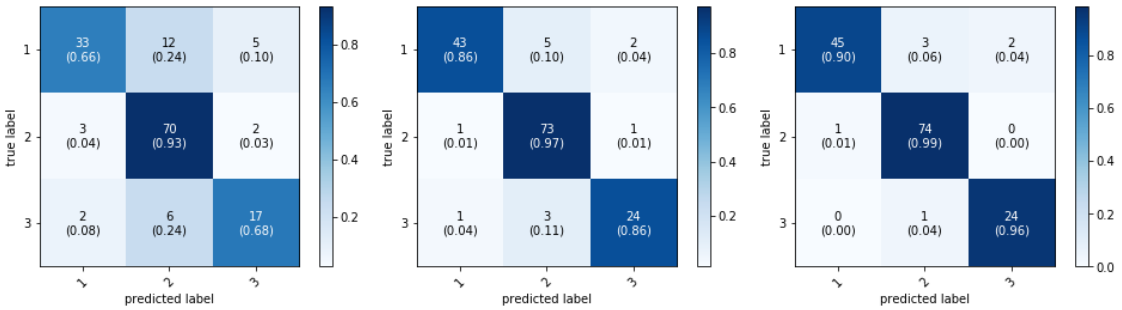
Based on the results, RF outperforms SVM for approximately all feature extraction models. Considering distinct models, DMD performed much better than the static models such as POD and PCA when using RF or SVM algorithms. With SVM, there is a jump in the accuracy level from POD with an averaged accuracy of 73% to PCA with an accuracy as high as 82% to DMD with the highest accuracy level of 92%. While with RF, these amounts are increased to 80%, 91% and 95% for POD, PCA and DMD, respectively. We can conclude that RF gives rise to higher levels of accuracy by shifting from static models to dynamic models in comparison with SVM. It means that a classifier trained with dynamic temporal features outperforms any of its static featured counterparts. This finding is consistently in agreement with the fact that liquid jet in crossflow has a very strong dynamic physics. The best classification prediction accuracy belongs to the DMD-RF model with the highest level of 95%, followed by the DMD-SVM and PCA-RF models with precision levels as high as 92% and 91%, respectively. Higher efficiency of DMD here is because the phase transformations between enhanced capillary, bag and multimode regimes are better captured by using temporal patterns. On the other hand, static models like POD and PCA lack any kind of temporal features and extract orthogonal SVD-based patterns. PCA also outperforms POD irrespective of the machine learning algorithm showing that the background subtraction

is effectively increasing the model accuracy. A smaller variance range is also remarkable for RF compared to SVM that emphasizes once again why the results extracted from RF are more reliable than SVM. It is also notable that most of the misclassified samples belong to the enhanced capillary breakup regime irrespective of the used model. The similarity between the trajectories of this breakup regime to other ones where the patterns are more complex and sparser in case of a high Weber number could be the reason for misclassification of samples in the enhanced capillary breakup regime. A visualized understanding of this fact could be observed in Figs.4.10-4.11 where the confusion matrices of different combined reduced-order and machine learning models are depicted. Breakup regimes of the enhanced capillary, bag and multimode are labelled with numbers 1,2 and 3, respectively. Both absolute and normal values are shown in each matrix element. These elements are either in the main diagonal or out of it, representing the average number of truly classified and misclassified samples, respectively.



(a) (b) (c)

Figure 4.10: Confusion matrix of different reduced-order models with SVM classifier, a)POD, b)PCA, and c)DMD



(a) (b) (c)

Figure 4.11: Confusion matrix of different reduced-order models with RF classifier, a)POD, b)PCA, and c)DMD

Despite both support vector machine and random forest classifiers' performance were satisfying in this study, extracted results should not be generalized to all other case studies. A voting block that decides the final class type for each sample based on the best three models of DMD-RF, DMD-SVM and PCA-RF is added to the source code of this paper. This section allows for better results where the number of classes, samples and other parameters change.

4.5 Conclusion

In summary, eigenmode-based feature extraction schemes are used in this study to classify image snapshots of liquid jets in crossflow. A better classification rate was obtained using random forest classifiers compared to support vector machines regardless of the used reduced-order model. This trend was the same while shifting from static models like proper orthogonal decomposition or principal component analysis to dynamic mode decomposition models irrespective of the machine learning algorithm. Afterwards, the best three predictive models were used to build a decision-maker block that returns the majority vote of them as the final answer. A graphical user interface application is also provided and is expected to be used for other case studies since the current built model is a generic one. Ultimately, the results of this work could be a step towards creating a closed-loop real-time controller that can update the input flow parameters to attain the desired breakup regime. There are various possible governing factors like the number of classes the effect of which are yet to be discovered in future works.

Chapter 5

Conclusions and Future work

In this chapter, the conclusion of this study is presented and the future perspectives in the area of data driven modelling of multiphase flow systems are proposed. It has to be mentioned out that the detailed results are summed up in the conclusion subsections of each chapter and the points noted here are the generalized outcomes of the work.

5.1 Conclusions

- In this work, we focused on liquid jets in crossflow and provided some approaches mostly based on the frequency spectra to interpret the extracted modes of POD and DMD methods.
- This interpretation was primarily demonstrated by analyzing some trivial cases like the spatio-temproal 2D wave and the von karman vortex street past a cylinder in the first chapter as well as the laminar jets of different inlet velocity in the second chapter.
- By making insights from the simple cases, we make an understanding of the fundamental spatial and temporal behaviour of the different regimes of liquid jets in crossflow.
- The effect of the aspect ratio is studied in the third chapter where different elliptic-type orifices are considered. It is concluded that the elliptical liquid

jets contain more coherent small-scaled structures with higher frequencies compared to the circular jets.

- In the final chapter, different eigenmode-based feature extraction schemes are utilized in order to classify the image snapshots of liquid jets in crossflow into various sub-regimes.
- Combinations of various reduced-order modelling techniques with machine learnt classifiers are considered and the most optimum joint model is presented.
- The results of this work could be a forward step towards creating a closed-loop real-time controller that can update the flow parameters to attain the desired breakup regimes.

5.2 Future work

- Data driven modelling could become an strong alternative in discovering new physical mechanisms and reviewing the existing laws from experimental and simulation fluid data. Using the recent developed methods like the sparse identification of nonlinear dynamics (SINDy) may be considered for further study of multiphase flow systems.
- Alongside the improvements in the area of reduced order modelling and data-driven optimization, the application of machine learning and deep learning to amend the performance and reduce the convergence time of solving complex problems seems crucial. Therefore, bringing into play modern machine learning environments as neural networks with different architectures is of interest.
- To be specific, Auto-encoders may be used as alternatives for the traditional reduced order modelling techniques like POD and DMD.
- We expect that the proposed discussions made in this work could be also exploited for analyzing more intricate multiphase flow systems where a non-newtonian liquid is present or there exists an effervescent nozzle or any other nozzle of different internal geometry, etc.

Bibliography

G. Amini and A. Dolatabadi. Axis-switching and breakup of low-speed elliptic liquid jets. *International Journal of Multiphase Flow*, 42:96–103, 2012.

M. Arienti and M. Soteriou. Time-resolved proper orthogonal decomposition of liquid jet dynamics. *Physics of Fluids*, 21(11):112104, 2009.

R. Arndt, D. Long, and M. Glauser. The proper orthogonal decomposition of pressure fluctuations surrounding a turbulent jet. *Journal of Fluid Mechanics*, 340:1–33, 1997.

S. Bagheri, P. Schlatter, P. Schmid, and D. Henningson. Global stability of a jet in crossflow. *Journal of Fluid Mechanics*, 624:33–44, 2009.

G. Berkooz, P. Holmes, and J. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25(1):539–575, 1993.

S. Bernero and H. Fiedler. Application of particle image velocimetry and proper orthogonal decomposition to the study of a jet in a counterflow. *Experiments in Fluids*, 29(1):S274–S281, 2000.

C. Brown, V. McDonell, and B. Kiel. Test bed for characterization of liquid jet injection phenomenon at augmentor conditions. In *42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, page 4569, 2006.

B. Brunton, L. Johnson, J. Ojemann, and N. Kutz. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *Journal of Neuroscience Methods*, 258:1–15, 2016.

S. Brunton and N. Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019.

S. Brunton, B. Noack, and P. Koumoutsakos. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.

T. Cambonie, N. Gautier, and J. Aider. Experimental study of counter-rotating vortex pair trajectories induced by a round jet in cross-flow at low velocity ratios. *Experiments in Fluids*, 54(3):1475, 2013.

F. Carvalho, M. Figueiredo, and A. Serpa. Flow pattern classification in liquid-gas flows using flow-induced vibration. *Experimental Thermal and Fluid Science*, 112: 109950, 2020.

S. Chakraborty and P. Das. Characterisation and classification of gas-liquid two-phase flow using conductivity probe and multiple optical sensors. *International Journal of Multiphase Flow*, 124:103193, 2020.

C. Chang and N. Dinh. Classification of machine learning frameworks for data-driven thermal fluid models. *International Journal of Thermal Sciences*, 135:559–579, 2019.

A. Chatterjee. An introduction to the proper orthogonal decomposition. *Current Science*, pages 808–817, 2000.

T. Colonius and K. Taira. A fast immersed boundary method using a nullspace approach and multi-domain far-field boundary conditions. *Computer Methods in Applied Mechanics and Engineering*, 197(25-28):2131–2146, 2008.

E. Curran. Scramjet engines: the first forty years. *Journal of Propulsion and Power*, 17(6):1138–1148, 2001.

Z. Dang, Y. Lv, Y. Li, and G. Wei. Improved dynamic mode decomposition and its application to fault diagnosis of rolling bearing. *Sensors*, 18(6):1972, 2018.

C. Duwig and P. Iudiciani. Extended proper orthogonal decomposition for analysis of unsteady flames. *Flow, turbulence and combustion*, 84(1):25, 2010.

A. Eroglu and R. Breidenthal. Structure, penetration, and mixing of pulsed jets in crossflow. *AIAA*, 39(3):417–423, 2001.

E. Farvardin. *Biodiesel spray characterization: a combined numerical and experimental analysis*. PhD thesis, Concordia University, 2013.

M. Farzad. Experimental study of rivulet/ice formation by colour-coded point projection method. Master's thesis, Concordia University, 2019.

K. Fujii and Y. Kawahara. Supervised dynamic mode decomposition via multitask learning. *Pattern Recognition Letters*, 122:7–13, 2019.

M. Gavish and D. Donoho. The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Transactions on Information Theory*, 60(8):5040–5053, 2014.

E. Gutmark and F. Grinstein. Flow control with noncircular jets. *Annual Review of Fluid Mechanics*, 31(1):239–272, 1999.

M. Herrmann. Detailed numerical simulations of the primary atomization of a turbulent liquid jet in crossflow. *Journal of Engineering for Gas Turbines and Power*, 132(6), 2010.

M. Herrmann, M. Arienti, and M. Soteriou. The impact of density ratio on the liquid core dynamics of a turbulent liquid jet injected into a crossflow. *Journal of Engineering for Gas Turbines and Power*, 133(6), 2011.

J. Higham, W. Brevis, and C. Keylock. Implications of the selection of a particular modal decomposition technique for the analysis of shallow flows. *Journal of Hydraulic Research*, 56(6):796–805, 2018.

S. Hirsh, K. Harris, N. Kutz, and B. Brunton. Centering data improves the dynamic mode decomposition. *arXiv preprint arXiv:1906.05973*, 2019.

G. Hobold and A. Da Silva. Machine learning classification of boiling regimes with low speed, direct and indirect visualization. *International Journal of Heat and Mass Transfer*, 125:1296–1309, 2018.

P. Holmes, J. Lumley, G. Berkooz, and C. Rowley. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press, 2012.

M. Jadidi, S. Moghtadernejad, and A. Dolatabadi. Penetration and breakup of liquid jet in transverse free air jet with application in suspension-solution thermal sprays. *Materials & Design*, 110:425–435, 2016.

M. Jadidi, V. Sreekumar, and A. Dolatabadi. Breakup of elliptical liquid jets in gaseous crossflows at low weber numbers. *Journal of Visualization*, 22(2):259–271, 2019.

Y. Kitamura and T. Takahashi. Stability of a liquid jet in air flow normal to the jet axis. *Journal of Chemical Engineering of Japan*, 9(4):282–286, 1976.

S. Krzeczkowski. Measurement of liquid droplet disintegration mechanisms. *International Journal of Multiphase Flow*, 6(3):227–239, 1980.

N. Kutz, S. Brunton, B. Brunton, and J. Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

S. Leask and V. McDonell. On the physical interpretation of proper orthogonal decomposition and dynamic mode decomposition for liquid injection. *arXiv preprint arXiv:1909.07576*, 2019.

Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio. Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision*, pages 319–345. Springer, 1999.

K. Lee, C. Aalburg, F. Diez, G. Faeth, and K. Sallam. Primary breakup of turbulent round liquid jets in uniform crossflows. *AIAA journal*, 45(8):1907–1916, 2007.

A. Liaw, M. Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

R. Madabhushi, M. Leong, M. Arienti, C. Brown, and V. McDonell. On the breakup regime map of liquid jet in crossflow. In *ILASS Americas, 19th Annual Conference on Liquid Atomization and Spray Systems, Toronto, Canada*, 2006.

M. Mancinelli, T. Pagliaroli, R. Camussi, and T. Castelain. On the hydrodynamic and acoustic nature of pressure proper orthogonal decomposition modes in the near field of a compressible jet. *Journal of Fluid Mechanics*, 836:998–1008, 2018.

J. Mann and N. Kutz. Dynamic mode decomposition for financial trading strategies. *Quantitative Finance*, 16(11):1643–1655, 2016.

R. Margason. Fifty years of jet in cross flow research. *Ceaj*, 1993.

M. Marzbali. *Penetration of circular and elliptical liquid jets into gaseous crossflow: a combined theoretical and numerical study*. PhD thesis, Concordia University, 2011.

J. Mazallon, Z. Dai, and G. Faeth. Primary breakup of nonturbulent round liquid jets in gas crossflows. *Atomization and Sprays*, 9(3), 1999.

K. Meyer, J. Pedersen, and O. Ozcan. A turbulent jet in crossflow analysed with proper orthogonal decomposition. *Journal of Fluid Mechanics*, 583:199–227, 2007.

S. Mittal and B. Kumar. Flow past a rotating cylinder. *Journal of Fluid Mechanics*, 476:303–334, 2003.

T. Murata, K. Fukami, and K. Fukagata. Nonlinear mode decomposition with convolutional neural networks for fluid dynamics. *Journal of Fluid Mechanics*, 882, 2020.

W. Noble. What is a support vector machine? *Nature Biotechnology*, 24(12):1565–1567, 2006.

B. Patte-Rouland, G. Lalizel, J. Moreau, and E. Rouland. Flow analysis of an annular jet by particle image velocimetry and proper orthogonal decomposition. *Measurement Science and Technology*, 12(9):1404, 2001.

K. Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2 (11):559–572, 1901.

E. Perlman, R. Burns, Y. Li, and C. Meneveau. Data exploration of turbulence simulations using a database cluster. In *Proceedings of the 2007 ACM/IEEE conference on Supercomputing*, pages 1–11, 2007.

A. Pollard, L. Castillo, L. Danaila, and M. Glauser. *Whither turbulence and big data in the 21st century?* Springer, 2016.

S. Prakash, A. Sinha, G. Tomar, and R. Ravikrishna. Liquid jet in crossflow–effect of liquid entry conditions. *Experimental Thermal and Fluid Science*, 93:45–56, 2018.

J. Proctor, S. Brunton, and N. Kutz. Dynamic mode decomposition with control. *SIAM Journal on Applied Dynamical Systems*, 15(1):142–161, 2016.

W. Rawat and Z. Wang. Deep convolutional neural networks for image classification: A comprehensive review, neural computing. *MIT Press Journals*, 29(9): 2352–2449, 2017.

C. Rowley. Model reduction for fluids, using balanced proper orthogonal decomposition. *International Journal of Bifurcation and Chaos*, 15(03):997–1013, 2005.

C. Rowley, I. Mezic, S. Bagheri, P. Schlatter, D. Henningson, et al. Spectral analysis of nonlinear flows. *Journal of Fluid Mechanics*, 641(1):115–127, 2009.

C. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. Henningson. Reduced-order models for flow control: balanced models and koopman modes. In *Seventh IUTAM Symposium on Laminar-Turbulent Transition*, pages 43–50. Springer, 2010.

K. Sallam, C. Aalburg, and G. Faeth. Breakup of round nonturbulent liquid jets in gaseous crossflow. *AIAA journal*, 42(12):2529–2540, 2004.

K. Sallam, C. Ng, R. Sankarakrishnan, C. Aalburg, and K. Lee. Breakup of turbulent and non-turbulent liquid jets in gaseous crossflows. In *44th AIAA Aerospace Sciences Meeting and Exhibit*, page 1517, 2006.

S. Sarkar, S. Ganguly, A. Dalal, P. Saha, and S. Chakraborty. Mixed convective flow stability of nanofluids past a square cylinder by dynamic mode decomposition. *International journal of heat and fluid flow*, 44:624–634, 2013.

P. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.

P. Schmid. Application of the dynamic mode decomposition to experimental data. *Experiments in Fluids*, 50(4):1123–1130, 2011.

K. Taira, S. Brunton, S. Dawson, C. Rowley, T. Colonius, B. McKeon, O. Schmidt, S. Gordeyev, V. Theofilis, and L. Ukeiley. Modal analysis of fluid flows: An overview. *Aiaa Journal*, 55(12):4013–4041, 2017.

G. Taylor. Generation of ripples by wind blowing over a viscous fluid. *The Scientific Papers of GI Taylor*, 3:244–254, 1940.

L. Trefethen and D. Bau. *Numerical linear algebra*, volume 50. Siam, 1997.

J. Tu and C. Rowley. On dynamic mode decomposition: theory and applications. *Journal of Computational Dynamics*, 1(2):391–421, 2014.

J. Tu, C. Rowley, D. Luchtenburg, S. Brunton, and N. Kutz. On dynamic mode decomposition: Theory and applications. *arXiv preprint arXiv:1312.0041*, 2013.

C. Weber. Disintegration of liquid jets. *Z. Angew. Math. Mech.*, 1:136–159, 1931.

P. Wu, K. Kirkendall, R. Fuller, and A. Nejad. Breakup processes of liquid jets in subsonic crossflows. *Journal of Propulsion and Power*, 13(1):64–73, 1997.

J. Xi and W. Zhao. Correlating exhaled aerosol images to small airway obstructive diseases: A study with dynamic mode decomposition and machine learning. *PloS one*, 14(1):e0211413, 2019.

G. Yunyi, L. Changwen, H. Yezhou, and P. Zhijun. An experimental study on droplet size characteristics and air entrainment of elliptic sprays. Technical report, SAE Technical Paper, 1998.

Appendix

Below is attached the python code for constructing the spatiotemporal signals.

```
# Importing Libraries
import numpy as np
import matplotlib.pyplot as plt

# Defining functions
def f1(x, t):
    return np.cos(x)*np.exp(1j*(1)*t)

def f2(x, t):
    return (np.tanh(x))*np.exp(1j*(2)*t)

# Setting the spatial and temporal boundaries
x = np.linspace(-2*np.pi, 2*np.pi, 30)
t = np.linspace(0, 5, 30)

# Defining the combined wave
X, T = np.meshgrid(x, t)
Z = f1(X, T)+f2(X,T)

# POD
U,S,VT = np.linalg.svd(Z, full_matrices=False)
V = VT.T
```

```
# DMD
# Finding the total number of images
total_images = np.size(t)
rank = total_images-1 #Arbitrary number in range (1,total_images-1)

# Defining the multiple matrices
X1 = np.array(X[:,0:total_images-1])
X2 = np.array(X[:,1:total_images])

# POD of X1
U,S,VT = np.linalg.svd(X1, full_matrices=False)
V = VT.T

# Constructing the reduced order SVD matrices
Ur = np.array(U[:,0:rank])
Sr = np.array(S[0:rank])
Vr = np.array(V[:,0:rank])

# Calculation of the linear operator A
Atilde = np.dot(Ur.T,np.dot(X2,np.dot(Vr,np.linalg.inv(np.diag(Sr)))))

# Eigenvalues & eigenvectors of the linear operator A
D, W = np.linalg.eig(Atilde)

# DMD modes calculation
Phi = np.dot(X2,np.dot(Vr,np.dot(np.linalg.inv(np.diag(Sr)), W)))

# Visualization of the modes (POD/DMD modes are denoted with U/Phi)
ind = np.arange(np.size(x))
plt.plot(ind,U[:,i],linewidth=1.5) # (i)= mode number
```