Machine Learning for Wavelet-based Valve Stiction Detection

Mr. Kris Prasopsanti

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Computer Engineering
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2022
Copyright of Chulalongkorn University

การเรียนรู้ของเครื่องสำหรับตรวจสอบปัญหาวาล์วฝืดด้วยสัญญาณเวฟเล็ต

นายกฤษฎิ์ ประสพสันติ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	Machine Learning for Wavelet-based Valve Stiction Detection
By	Mr. Kris Prasopsanti
Field of Study	Computer Engineering
Thesis Advisor	Assistant Professor SUKREE SINTHUPINYO, Ph.D.

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

.... Dean of the FACULTY OF ENGINEERING
(Professor SUPOT TEACHAVORASINSKUN, D.Eng.)

THESIS COMMITTEE

.... Chairman
(Assistant Professor NATTEE NIPARNAN, Ph.D.)
.... Thesis Advisor
(Assistant Professor SUKREE SINTHUPINYO, Ph.D.)
.... Examiner
(JESSADA THUTKAWKORAPIN, Ph.D.)
.... External Examiner
(Assistant Professor Denduang Pradubsuwun, D.Eng.)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

กฤษฎิ์ ประสพสันติ : การเรียนรู้ของเครื่องสำหรับตรวจสอบปัญหาวาล์วฝืดด้วยสัญญาณ
เวฟเล็ต. (Machine Learning for Wavelet-based Valve Stiction Detection) อ.ที่
ปรึกษาหลัก : ผศ. ดร.สุกรี สินธุภิญโญ

ปัญหาวาล์วฝืด (valve stiction) คือปัญหาสำคัญในอุตสาหกรรมเนื่องจากส่งผลให้เกิดการสั่นและการขัดขวางในการควบคุมอัตราการไหลของสายการผลิต งานวิจัยนี้ได้ศึกษาการตรวจจับการติดขัดของวาล์วโดยการวิเคราะห์ข้อจำกัดของวิธีการที่มีอยู่ในปัจจุบัน และเสนอวิธีการตรวจจับใหม่ซึ่งผสมการสร้างใหม่เชิงเวฟเล็ต (wavelet reconstruction) และโครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Networks: CNN) เพื่อเพิ่มประสิทธิภาพในการตรวจจับ วิธีการนี้ใช้รูปแบบของกราฟความสัมพันธ์ระหว่างค่าตัวแปรในกระบวนการ (Process Variable: PV) และคำสั่งจากตัวควบคุม (controller output: OP) ซึ่งถูกปรับปรุงให้แสดงลักษณะเฉพาะของการติดขัดอย่างชัดเจน เป็นข้อมูลป้อนเข้าสู่แบบจำลอง CNN โดยข้อมูลที่ใช้ในการฝึกฝนแบบจำลองเป็นข้อมูลจำลองและข้อมูลที่ใช้ในการทดสอบเป็นข้อมูลจริงจาก International Stiction Data Base (ISDB) โดยใช้ F1 score เป็นตัววัดประสิทธิภาพหลัก ผลการทดลองแสดงว่าการปรับปรุงข้อมูลป้อนเข้า PV(OP) ร่วมกับ CNN สามารถทำให้ได้ F1 score สูงสุดที่ 0.9 ซึ่งสูงกว่าวิธีการที่ใช้อยู่ในปัจจุบัน และการวิเคราะห์รายละเอียดของกรณีที่มีความผิดพลาดในการตรวจจับ ทำให้ได้ความเข้าใจเกี่ยวกับจุดเด่นและข้อจำกัดของวิธีการนี้ เช่น ความสามารถในการตีความผลลัพธ์ (interpretability) และความทนทานของแบบจำลอง (robustness) การตรวจจับการติดขัดของวาล์วอย่างแม่นยำนี้จะช่วยให้เราสามารถวางแผนการบำรุงรักษาและการปรับปรุงอย่างมีประสิทธิภาพ ทำให้กระบวนการอุตสาหกรรมมีความปลอดภัยและประสิทธิผลเพิ่มขึ้น

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต ....
ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

6372003221 : MAJOR COMPUTER ENGINEERING

KEYWORD: Valve stiction detection, Convolutional neural networks, Wavelet reconstruction

Kris Prasopsanti : Machine Learning for Wavelet-based Valve Stiction Detection. Advisor: Asst. Prof. SUKREE SINTHUPINYO, Ph.D.

Valve stiction presents challenges in industrial process control, leading to oscillations and hindering the regulation of fluid flow. This thesis addresses the detection of valve stiction by exploring the limitations of current methods and proposes a novel approach that combines wavelet reconstruction and convolutional neural networks (CNN) to enhance stiction detection performance. The proposed method utilizes preprocessed process variable versus controller output (PV(OP)) plots as input to the CNN model, capitalizing on the distinctive characteristics of stiction. Training and evaluation employ both simulated and real-world data from the International Stiction Data Base (ISDB), with the F1 score serving as the primary performance metric. Results demonstrate that the modified PV(OP) input approach, in conjunction with the CNN classifier, achieves an impressive F1 score of 0.9, surpassing conventional methods. In-depth analysis of fault cases provides valuable insights into the strengths and limitations of the approach, emphasizing interpretability and robustness. The accurate detection of valve stiction enables proactive maintenance and targeted interventions, ultimately improving the performance of interconnected systems and enhancing safety and efficiency in industrial processes.

Field of Study:	Computer Engineering	Student's Signature ....
Academic Year:	2022	Advisor's Signature ....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Asst. Prof. Sukree Sinthupinyo, my advisor, for their guidance and support throughout this thesis. His valuable input and advice have been essential to the successful completion of this research. I am also thankful to the faculty members of the computer and chemical engineering departments Chulalongkorn university for their contributions and knowledge sharing, which have greatly enriched this study. Lastly, I want to extend my appreciation to my family and friends for their love, encouragement, and support. Their presence has been a source of strength and motivation throughout this academic journey.

Kris Prasopsanti

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

TABLE OF CONTENTS

Page

.... iii

ABSTRACT (THAI) iii

.... iv

ABSTRACT (ENGLISH) iv

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER I INTRODUCTION 11

  1.1 Background and motivation 11

  1.2 Research objective 14

  1.3 Scopes of work 14

  1.4 Expected benefits 15

CHAPTER II RELATED THEORY 17

  2.1 Industrial process control loop fundamental 17

  2.2 Valve stiction behavior 19

  2.3 Technique used in this research 20

    2.3.1 Power spectral density 20

    2.3.2 Butterworth filter 21

    2.3.3 Wavelet decomposition 21

    2.3.4 Wavelet reconstruction 22

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.3.5 Artificial neural networks and convolutional neural network 23
CHAPTER III LITERATURE REVIEW.. 25
3.1 Valve stiction model.. 25
3.2 Traditional detection methods.. 28
3.3 Machine learning detection methods.. 33
CHAPTER IV METHODOLOGY 39
4.1 Data generation for model training 39
4.2 Data preprocessing... 40
4.3 Convolutional neural network model architecture 41
4.4 Test set and evaluation 42
CHAPTER V RESULT AND DISCUSSION.. 45
5.1 Model input interpretability... 45
5.2 Performance evaluation.. 47
5.3 Result comparison 51
5.4 Fault cases analysis... 52
5.5 Sensitivity analysis... 56
5.6 Other neural networks 61
CHAPTER VI CONCLUSION... 63
APPENDIX A 66
DATA GENERATION IN LITERATURES 66
REFERENCES 71
VITA... 74

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF TABLES

Page

Table 1. Comparison of classical detection methods 33

Table 2. Structure of Kamaruddin BSN CNN model 36

Table 3. Structure of Zhang CNN models 36

Table 4. Comparison of machine learning detection methods 38

Table 5. Parameters of data generation 40

Table 6. Test dataset 44

Table 7. Comparision of F1 Score with different models and inputs.............................. 52

Table 8. Fault cases of different methods 54

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF FIGURES

Page

Figure 1 Control loop with a valve as an actuator 19

Figure 2. Choudhury valve signature of stiction process 20

Figure 3. Flow chart of Choudhury model 26

Figure 4. Flow chart of Kano model 27

Figure 5. He valve signature of stiction process 28

Figure 6. Flow chart of He model 28

Figure 7. A normalized index for stiction and aggressive control cases 30

Figure 8. Distributions of second derivative of process output in 31

Figure 9. Examples of the data in ISDB 32

Figure 10. 8x8 pixels input for Dambos model 35

Figure 11. Input shape of BSD method 35

Figure 12. Convolutional neural network model architecture 42

Figure 13. PV(OP) Plots obtained from both non-strict and strict cases 46

Figure 14. Confusion matrix 47

Figure 15. Model input examples for feature map visualization 48

Figure 16. Second convolutional layers output of Stiction 1 example 48

Figure 17. Second convolutional layers output of Stiction 2 example 49

Figure 18. Second convolutional layers output of Non-Stiction 1 example 49

Figure 19. Second convolutional layers output of Non-Stiction 2 example 50

Figure 20. ROC curve 51

Figure 21. Model performance with different number of epochs 57

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 22. Model performance with different number of output filters in convolutional layers 57

Figure 23. Model performance with different leaning rate 58

Figure 24. Model performance with different wavelet reconstruction threshold 59

Figure 25. Model performance distribution with 20 different training sets 59

Figure 26. Model performance distribution with process time constant = 0.1 minutes for10 different training sets 60

Figure 27. Model performance distribution with process time constant = 0.4 minutes for10 different training sets 60

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER I

INTRODUCTION

1.1 Background and motivation

Valve stiction is one of the most prominent problems in chemical process monitoring. In each chemical plant there are 500 – 5,000 process control loops, most of them have fluid as their main element and need valve as the primary control equipment. As those valves tend to have irregular maintenance, they stick regularly. According to Raul "Stiction is a combination of the words stick and friction, created to emphasize the difference between static and dynamic friction. Stiction exists when the static (starting) friction exceeds the dynamic (moving) friction inside the valve. Stiction describes the valve's stem (or shaft) sticking when small changes are attempted." [1]. In other words, it is an occurrence that prevents valves from moving in a manner that a plant operator or controller wants because of forces of friction.

While the problem itself is not a big deal, when combined with how typical controllers in the industries work, it causes limited oscillations in the fluid flowrate as well as all related state variables of the process. This oscillation is a major concern in industries as about 80% of their closed loop have the problem [2] resulting in excessive energy and raw materials consumption as well as safety issues. However, even though several surveys showed that valve stiction is a cause of around 20-30% of the oscillation problem [3-5], it is not the only one. The oscillation problem can be caused by others such as external disturbances, inadequate controller tuning or presence of other process non-linearities. And while the solutions of oscillation problems are relatively straight-forward if the causes are known, identifying the causes can be problematic.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Existing methods for valve stiction detection fall into two categories. The first category comprises traditional feature engineering-based approaches, which rely on engineering principles and domain knowledge to identify stiction behavior from process data. While these methods offer interpretability, they are limited in their applicability to real-world industrial data and often require detailed system knowledge. For instance, the Bicoherence and Ellipse-fitting method (BIC), introduced by Choudhury et al. in 2006 [6], is a prominent example and one of the best performed methods in this category. BIC identifies stiction behavior based on Gaussian and linearity tests, as well as the ellipse shape of the frequency-filtered Controller Output-Process Variable (PV(OP)) plot. However, these tests require long lengths of well-behaved data, which can restrict their usability.

In recent years, there has been a growing interest in utilizing deep learning approaches for valve stiction detection. However, the application of data science techniques in the chemical industry is not without challenges. One major hurdle is the scarcity of labeled data for analysis and model training. Obtaining labeled data from real-world industrial processes can be difficult and time-consuming. As a result, many researchers have resorted to using simulation data as a substitute. Meaning, the sensitivity of deep learning models to how the training set is simulated is a significant concern. Simulation parameters and assumptions can significantly impact the performance of the model, making it challenging to ensure robustness and applicability in practical industrial settings.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Another challenge stems from the black-box nature of deep learning methods. While these models can achieve impressive performance, their inner workings are often difficult to interpret. This lack of interpretability can hinder the ability to gain insights into the underlying mechanisms of valve stiction and the specific patterns driving its occurrence. Additionally, the reliance on small test sets during model evaluation can introduce a level of uncertainty regarding the generalizability of the results to real-world scenarios.

To address these issues, we have integrated traditional methods and fundamental process engineering knowledge into our work to enhance interpretability and reduce the dependence on simulating the training set. Specifically, we employ the main idea of Ellipse-fitting approach to transform simulated data and real-world industrial data to make them more similar to each other, regardless of the simulation parameters used. However, for the frequency filtering method, we only filter the signal with a frequency lower than the oscillation frequency of the data to retain as much information as possible, which makes the problematic Gaussian and linearity tests unnecessary. As for the high frequency components, inspired by Xu et al. valve stiction detection utilizing wavelet technology [7], we use wavelet reconstruction to further de-noise and prepare a suitable PV(OP) plot for training and testing the Convolutional Neural Network (CNN) model. The main selling point of our idea is, despite the additional steps involved, our valve stiction detection process remains automated, harnessing the advantages of deep learning methods in terms of time efficiency and minimal system knowledge requirements. By incorporating these methods and knowledge, we believe that the practicality and robustness of the deep learning approach can be significantly enhanced.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

For the model CNN architecture is well-suited for the task of valve stiction detection in chemical process control. CNNs have demonstrated remarkable success in various domains, particularly in image and signal processing tasks. This is because CNNs excel at capturing and extracting meaningful patterns and features from high-dimensional data. In the context of valve stiction detection, the PV(OP) plot can be considered as a form of signal data that encapsulates crucial information about stiction behavior. By leveraging the inherent properties of CNNs the network can effectively analyze the patterns present in the PV(OP) plot.

In conclusion, our work represents a novel integration of traditional methods and data science techniques to address the challenges of valve stiction detection in chemical process control. By combining the interpretability of traditional approaches and the power of deep learning, we aim to develop a reliable and practical solution for detecting and mitigating valve stiction. We anticipate that our approach will contribute to the enhancement of industrial processes, promoting energy efficiency, reducing material waste, and ensuring process safety.

1.2 Research objective

Develop a new convolutional neural network-based valve stiction detection method with comparable performance with current research and is ready to employ in real-world environment.

1.3 Scopes of work

Review and analyze existing traditional feature engineering-based methods for valve stiction detection including their limitations and applicability to real industrial data.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Explore the application of deep learning approaches for valve stiction detection and identify the challenges associated with employing data science techniques in the chemical industry.

Develop an integrated approach that combines traditional methods, process engineering knowledge, and deep learning techniques to enhance interpretability, reduce reliance on simulated training sets, and improve the accuracy and adaptability of valve stiction detection.

Implement a frequency filtering method that focuses on retaining valuable information eliminating the need for Gaussian and linearity tests and prepare suitable PV(OP) plots for training and testing a Convolutional Neural Network (CNN) model.

Evaluate and analyst the proposed approach using real-world industrial data from International Stiction Data Base (ISDB) benchmark.

1.4 Expected benefits

Improved process efficiency: By accurately detecting and mitigating valve stiction, the proposed method can significantly enhance process efficiency and improve overall productivity in industrial processes.

Enhanced process safety: Stiction-induced oscillations in fluid flow can compromise process stability and safety. By detecting and addressing valve stiction, the proposed method contributes to maintaining safe operating conditions, reducing the risk of equipment failure, and preventing accidents in industrial settings.

Automation and efficiency: The integration of deep learning techniques automates the stiction detection process, reducing the need for manual intervention. This improves the efficiency of monitoring and control systems, enabling faster and more accurate identification of stiction issues.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Advancement of knowledge: The research contributes to the existing body of knowledge by integrating traditional methods with data science techniques. It demonstrates the potential of combining interpretability with the power of deep learning algorithms, fostering advancements in the field of valve stiction detection and process control systems in general.

Future research opportunities: The proposed method opens avenues for further research and exploration. Future studies can investigate the application of alternative neural network architectures, explore the integration of additional features or data sources, or refine the proposed method to address specific challenges in valve stiction detection and control loop optimization.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER II

RELATED THEORY

In this chapter, the related theory is organized into three sections. Firstly, the Industrial process control loop fundamentals discusses the basic principles of process control that are necessary to understand valve stiction behavior. Secondly, the chapter explores valve stiction behavior itself. Finally, the techniques employed in this research are discussed.

2.1 Industrial process control loop fundamental

The problem of valve stiction is happened by an interactive between a sticky valve and process system controller, so in order to understand the problem one needs to know the basic of the control loop, specifically nature of single input single output closed loop feedback control (SISO). In industrial process control, control loops play a vital role in maintaining the desired performance and stability of various processes. These control loops are widely used in industries such as chemical, oil and gas, power generation, and manufacturing. Understanding the fundamental components and concepts of control loops is essential for comprehending the challenges associated with valve stiction detection.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

A basic control loop consists of four key components: a process or system, a measurement device, a controller, and a final control element. The process or system represents the physical system being controlled, which could be a chemical reactor, a distillation column, or any other industrial process. The measurement device is responsible for monitoring the process variable (PV), which is the output or characteristic being controlled, such as temperature, pressure, or flow rate. The controller receives the PV from the measurement device and compares it with the setpoint (SP), which is the desired value for the PV. Based on this comparison, the controller calculates the appropriate control action (OP) to be applied.

The control action is then transmitted to the final control element, which is typically a valve responsible for adjusting the flow of a fluid or altering the input to the process. The valve output (MV) modifies the process conditions to bring the PV closer to the SP. This feedback loop continues, with the measurement device continuously providing information to the controller for ongoing adjustment.

To go into greater depth, the main idea of process control is to control PV in a state-determined system to be as close to SP as possible. Where state variables are a non-unique minimum set of variables that fully describe the system and the state-determined system model is a model which has a characteristic that a mathematical description of the system in terms of state variables x_i together with knowledge of those variables at an initial time t_0 and the system inputs for time $t > t_0$, are sufficient to predict the future system state and outputs for all time $t > t_0$. In other words, ignoring process time delay, history of the system states before the time of interest do not have any effect on how to control the process in the future.

With this characteristic, theoretically any PV can be controlled by manipulating system inputs, which in chemical industries usually can be achieved by changing some fluid flowrate by adjust valve output. Typical control structure is shown in Figure 1. Firstly, current PV is compared with SP then the difference is sent to a controller which will calculate and give OP to change MV in a direction that reduces the difference. There are many types of controllers in industries, but the most commonly used ones are some variances of proportional integral derivative controllers (PID) which calculates OP based on eq. 1. As the name suggests, OP of PID is calculated with a combination of 3 elements. Firstly, a proportional element (P) which adjusts the controller output to reduce error between SP and PV by change the output proportionally to the error, then an integral element (I) which add an integrator to a system to eliminate PV offset and, lastly, derivative element (D) to deal with process with high frequency setpoint changes. In practice however,

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

typically only 1 or 2 of these elements are applied. This is because, even though, derivative element improves performance and speed of the controller, it amplifies signal will high fundamental frequency making process more prone to noises. And ideally, we need exactly 1 integrator in a process, so in normal self-regulated processes we will use PI controller but for ramp processes which already has an integrator we will use only P controller.

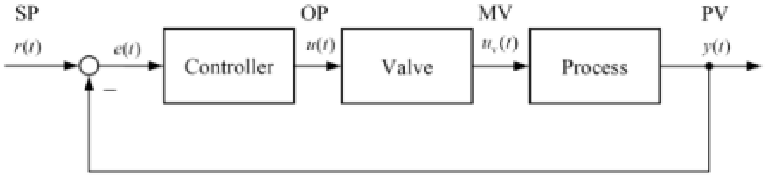


Figure 1 Control loop with a valve as an actuator

$$u(t) = K(e(t) + I\int e(t)dt + D\frac{de(t)}{dt}) \quad (1)$$

2.2 Valve stiction behavior

In closed control loop system, valve stiction is an occurrence that prevents valve to move as a process controller command due to force of frictions. As of common knowledge, friction forces can be classified to static and dynamic forces which affect stationary and moving objects respectively. For a standard control loop as shown in Figure 1, as a controller directs valve to change its output it will ideally change accordingly resulted in a straight 45-degree line from point A as shown in valve signature in Figure 2. However, in the presence of valve stiction, the valve will not move as a driving force from controller is not enough to overcome its static friction force. This will force the controller to increase its output until it is higher than the friction. Then as the valve starts moving at point B, friction will change to dynamic force which is lower than its static counterpart. So, the driving force will be higher than friction force resulting in valve jump to point C and slip to point D. This jump and slip will cause valve output to change more than the controller intention

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

to point G forcing controller to command in the opposite direction and the cycle continues. This is the cause of a limited oscillation in fluid flowrate which can be represented by a quadrilateral BGHM [8].

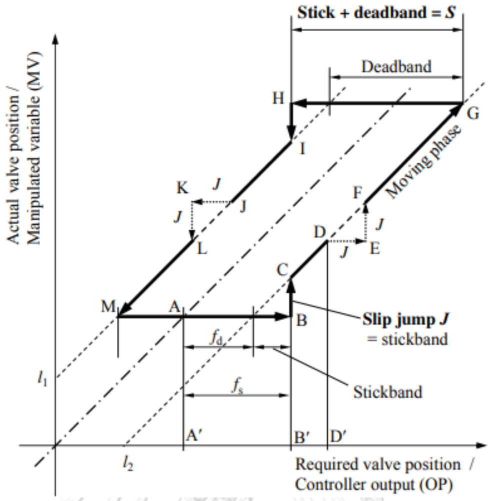


Figure 2. Choudhury valve signature of stiction process

2.3 Technique used in this research

2.3.1 Power spectral density

Power spectral density (PSD, S_x) is a fundamental concept in signal processing that characterizes the distribution of power across different frequencies in a signal. It provides valuable insights into the frequency content and distribution of power in a signal, allowing for analysis of its spectral properties.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

PSD is computed by taking the Fourier transform of the auto-covariance function (ACF, r_{xx}) as shown in eq. 2 and eq. 3 leveraging the idea that ACF can preserve oscillation compared to raw data. Fourier transform converts a signal from the time domain to the frequency domain, representing the signal as a sum of sinusoidal components at different frequencies. Squaring the magnitude of the spectrum gives the power associated with each frequency component [9]. It is a valuable tool in signal processing for analyzing the power distribution across different frequencies in a signal. It provides insights into the spectral characteristics and frequency content of a signal, aiding in various tasks such as signal analysis, filtering, and feature extraction.

$$S_x(f) = \sum_{k=-\infty}^{\infty} r_{xx}(k) e^{-2\pi jfk} \tag{2}$$

$$r_{xx}(k) = \frac{\sum_{t=1}^{N-k} (x(t) - \overline{x})(x(t+k) - \overline{x})}{\sum_{t=1}^{N} (x(t) - \overline{x})^2} \tag{3}$$

2.3.2 Butterworth filter

The Butterworth filter is a type of electronic filter used in signal processing to attenuate unwanted noise or extract specific frequency components from a signal. It is a type of band-pass filter [10]. The filter's key characteristic is its frequency response, which exhibits a smooth and monotonic passband and a steep roll-off in the stopband. The filter design employs polynomial approximation, resulting in a maximally flat passband and a Butterworth response.

The filter's order corresponds to the number of poles in its transfer function and affects the steepness of the roll-off in the stopband. Higher-order filters exhibit more pronounced roll-offs but are more intricate to implement. The Butterworth filter finds widespread application in audio processing, speech processing, biomedical signal processing, and various other domains.

2.3.3 Wavelet decomposition

Wavelet decomposition is a powerful technique used in signal and image processing to analyze and extract information at different scales. It involves breaking down a signal or an image into its constituent parts, revealing both global and local features. This process utilizes wavelet functions that are well-localized in both the time and frequency domains as shown in eq. 4, where a, b and φ are scale, location, and basis wavelet respectively. By convolving these wavelet functions with the signal or image, a set of coefficients is obtained, representing the strength of the wavelet at each location and scale [11].

$$T(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}x(t)\varphi\frac{t-b}{a} \tag{4}$$

The wavelet decomposition process starts with the original signal or image and proceeds in a hierarchical manner. At each level of decomposition, the signal or image is divided into approximation and detail coefficients. The approximation coefficients capture the low-frequency components, representing the overall trends or coarse details, while the detail coefficients represent the high-frequency components, capturing the finer details or rapid variations.

To obtain these coefficients, the wavelet functions are convolved with the signal or image. The convolution involves multiplying the wavelet function with the localized portion of the signal or image and summing the results. This operation is performed at different positions and scales to compute the coefficients at each level of decomposition.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

2.3.4 Wavelet reconstruction

Wavelet reconstruction is the process of synthesizing a signal or an image from its wavelet coefficients obtained through wavelet decomposition. It is the inverse operation of wavelet decomposition and allows for the reconstruction of the original data. The reconstruction process starts with the finest level of decomposition and proceeds in a reverse hierarchical manner [12].

Wavelet reconstruction finds applications in various fields, such as image compression, denoising, and signal analysis. It provides a flexible and efficient approach to reconstructing signals and images with high accuracy while preserving important features. The hierarchical nature of wavelet reconstruction allows for a multi-resolution analysis, enabling the reconstruction of data at different levels of detail and facilitating the extraction of valuable information.

2.3.5 Artificial neural networks and convolutional neural network

Artificial neural networks (ANNs) and Convolutional neural networks (CNNs) are two popular types of deep learning models used for various machine learning tasks. Both ANNs and CNNs are inspired by the structure and function of the human brain, where complex information processing and pattern recognition occur.

An Artificial Neural Network (ANN) is a computational model composed of interconnected artificial neurons, also known as nodes or units. These nodes are organized into layers, typically consisting of an input layer, one or more hidden layers, and an output layer. Each node receives input signals, applies an activation function to them, and produces an output signal that is propagated to the next layer. The strength of the connections between nodes, known as weights, is adjusted during training to optimize the network's performance. ANNs are widely used for various tasks, including classification, regression, and pattern recognition.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

On the other hand, a Convolutional Neural Network (CNN) is a specialized type of neural network designed specifically for processing structured grid-like data, such as images. CNNs excel in image analysis tasks due to their ability to automatically learn and extract hierarchical features from input data. Unlike ANNs, CNNs are characterized by their unique layers, including convolutional layers, pooling layers, and fully connected layers.

The convolutional layers in a CNN consist of a set of learnable filters that convolve over the input image, performing local operations to extract features.

These filters scan the image using a sliding window approach, capturing different patterns and detecting features at different scales. The pooling layers, such as max pooling or average pooling, reduce the dimensions of the feature maps while preserving important information. Finally, fully connected layers at the end of the network aggregate the extracted features and make predictions based on them.

The key advantage of CNNs lies in their ability to automatically learn spatial hierarchies of features from raw data. By leveraging the local connectivity and weight sharing of convolutional layers, CNNs can capture important patterns and structures in images, making them highly effective for tasks such as image classification, object detection, and image segmentation.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER III

LITERATURE REVIEW

In this chapter, the literature is divided into three sections to provide a comprehensive understanding of the topic. Firstly, the development of various valve stiction models is explored, aiming to accurately describe and simulate this phenomenon. Secondly, the chapter delves into the traditional methods that are currently employed in industries to address valve stiction. Finally, the chapter investigates the emerging trend of utilizing deep learning approaches as a potential solution to tackle this issue.

3.1 Valve stiction model

The aforementioned concept mentioned in the theory section is the core idea of the most widely accepted valve stiction models with 2 parameters, Stick + Deadband (S) and Slip jump (J), proposed by Choudhury [8] and Kano [13] which represented by flow charts shown in Figure 3 and 4 respectively. From the flowcharts, firstly both check whether the controller output is within 0 to 100% range, then calculate change in the output and compare with previous change. If the changes do not have the same sign or the valve stop moving, the valve got stuck until difference between controller output and current output is higher than S. On the other hand, if the changes have the same sign and the valve is moving, valve moves by the controller output minus dynamic fiction, (S-J)/2.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

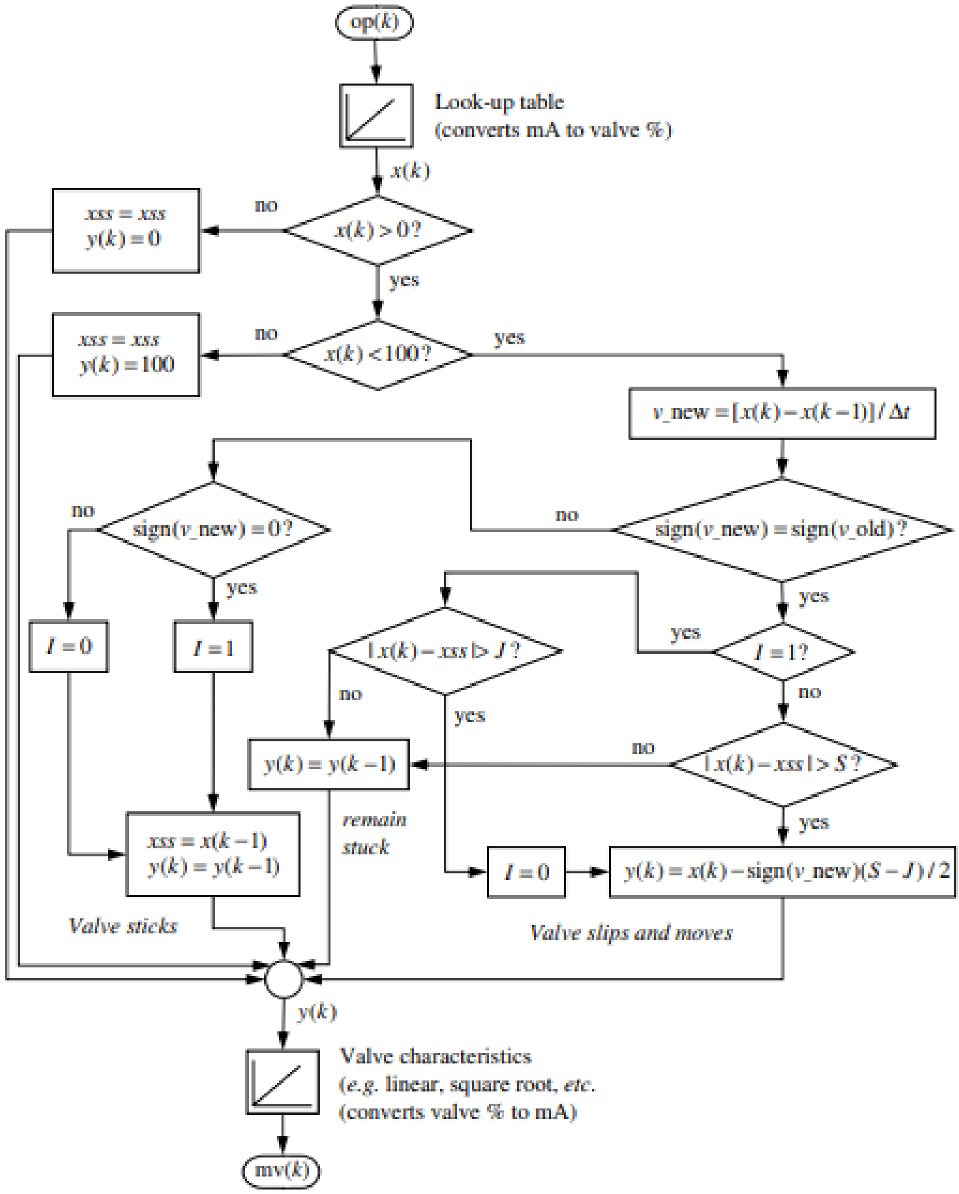


Figure 3. Flow chart of Choudhury model

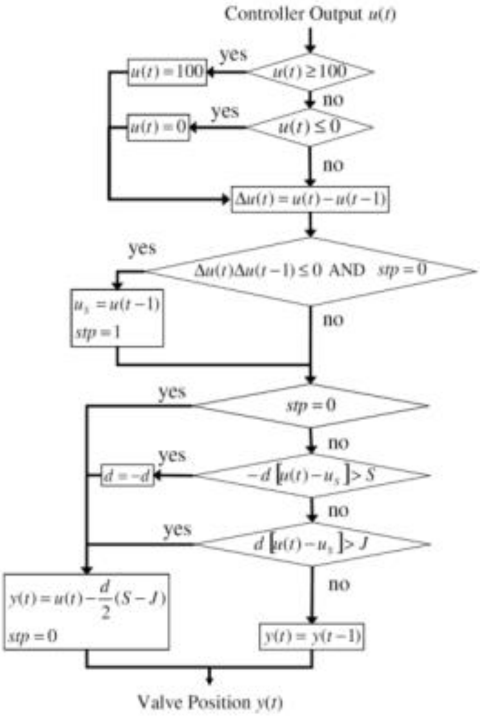


Figure 4. Flow chart of Kano model

Nevertheless, these models assumed that that valve only stops moving when the controller output changes direction. this is not true in cases of severe stiction which leads to another widely used model was proposed by He et al. [14]. In severe cases, He showed that, as valve dynamic is vastly faster than rate of change in controller output, the valve always stops before next controller direction comes. They also suggested that, due to the second order nature of valve dynamic, every time the valve moves, it will move with an overshoot around 0.99 of the driving force. Those suggestions led to a valve signature in Figure 5 and flowchart in Figure 6. Nevertheless, it is important to note that those models are not entirely accurate representation of valve stiction characteristics and the behavior in real world can be unpredictable and stochastic. While He model may be better in cases of high stiction, it will give an inaccurate representation if the stiction is not severe.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

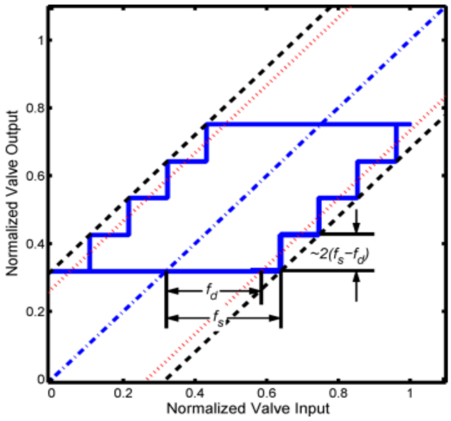


Figure 5. He valve signature of stiction process

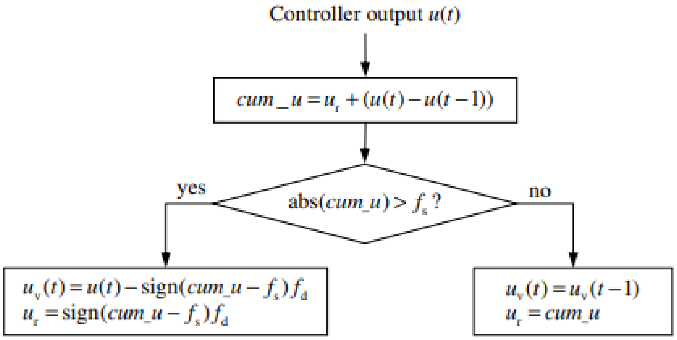

Controller output $u(t)$
$cum\_u = u_r + (u(t) - u(t-1))$
yes
$\text{abs}(cum\_u) > f_s$?
no
$u_v(t) = u(t) - \text{sign}(cum\_u - f_s)f_d$
$u_r = \text{sign}(cum\_u - f_s)f_d$
$u_v(t) = u_v(t-1)$
$u_r = cum\_u$

Figure 6. Flow chart of He model

3.2 Traditional detection methods

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Several detection methods have been developed for valve stiction detection over the past decades. While this work may focus on how to improve the machine learning approaches, many traditional concepts were used as a fundamental to design the machine learning pipelines. The most prominent traditional detection methods can be classified into 4 groups: PV(OP) plot, Wave shape of OP or PV, Cross-correlation and non-linear stiction model.

The PV(OP) plot is the most well-established stiction detection method. The main idea is, as shown in Figure 2, as we plot a relationship between controller output (OP) and valve output (MV) of the stiction control loop, we will get a quadrilateral graph. However, in real-world valve output data is typically not available, so relationship between OP and process variable (PV) is used instead. The most well-known method in this category is Bicoherence and Ellipse-fitting method (BIC) introduced by Choudhury et al [2]. In this method, first, data is tested by bispectrum and bicoherence for non-Gaussian and non-linearity respectively. As both have been confirmed data are filtered by Wiener filter to remove their moving trend. Then in the case of stiction PV(OP) plot of filtered data should show elliptic pattern and stiction degree, S, can be determined by maximum width of the ellipse along OP axis. However, while this method has high performance, bispectrum and bicoherence require at least 1,024 (preferably 4,096) data points without any step change or abrupt change to be reliable making this method impractical for many real-world scenarios.

As for wave shape of OP and PV, Area-peak method (AREA) proposed by Salsbury and Singhal) [15], Curve-fitting method (CURVE) by He et al. [14] and Relay method (RELAY) by Rossi and Scali [16] are the most widespread ones. In Area-peak method, characteristic of the process is determined by shape of the PV signal between zero-crossing events, a normalized index, R is defined as ratio between area before and after the peak as shown in Figure 7 and it claims that, in cases of stiction, the R value will be significantly higher than one. Then in curve-fitting method, the idea is a first integrated signal after valve output, which is OP in case of self-regulated process and PV in case of ramp process, should be in a triangular shape in stiction cases, whereas the signal should be in sine shape if the oscillation is caused by any other means. At last, the Relay method is a variation of the Curve-fitting method, but instead of only 2 mentioned shapes, this method also tries fitting a shape of a relay system and use one that has better approximation between the relay and the triangular shape to calculate stiction parameters.

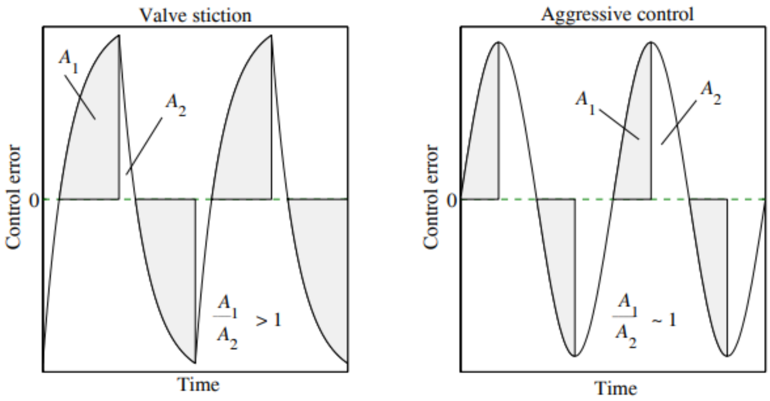
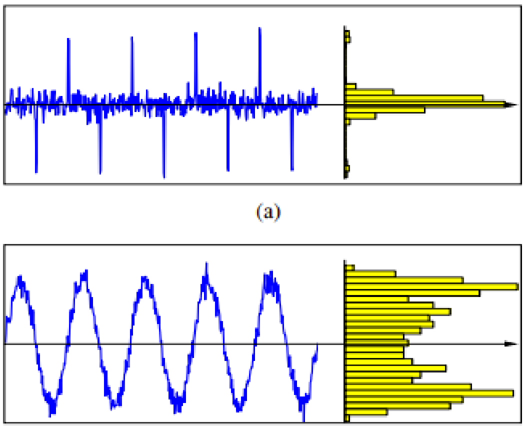


Figure 7. A normalized index for stiction and aggressive control cases

Another feature engineering-based detection method discussed in this study is Xu's idea of utilizing wavelet technology for wave shape-based striction detection. In 2009, Xu showed that due to Lipschitz regularity theory, the jump of the PV signal from valve stiction can be differentiated from the jump from noise via wavelet reconstruction. As the magnitude of the noise wavelet coefficients decays as scales increase, it gets filtered out in the reconstruction process, while the magnitude of the valve stiction wavelet coefficients remain the same. Therefore, they can detect the stiction behavior by looking at a jump after the reconstruction.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

For a Cross-correlation-based detection (CORR) (Horch et al.) [17], the crux of the method is that for self-regulation processes, in cases of valve stiction, OP and PV wave should have phase lag ~ Π whereas in other cases such as external disturbances the phase lag is ~ 2Π. So, if cross-correlation between OP and PV is an odd function, there is stiction problem otherwise, the oscillation is happened by other causes. However, this method requires the presence of PI controller, so it cannot be used in ramp control loop. In case of ramp processes, Hoch proposed a histogram detection method (HIS). The main idea of the method is, for striction process, distribution of second derivative of the process output will only have 1 peak while the second derivative of process output oscillated by other causes will has 2 peaks as shown in Figure 8.


(a)

(b)

Figure 8. Distributions of second derivative of process output in (a) stiction, (b) other cases

Lastly, for of non-linear stiction model, Hammerstein-model-based Methods (Jelali [18], Lee el al. [19], Karra and Karim [20]) are used to represented characteristics of stiction processes. In addition of process model, Hammerstein structures also add some form of stiction model discussed in previous sector. Then, in order to connect S and J to physical values, they are represented by static (f_s) and dynamic friction (f_d) as a relationship shown in eq. 5 and 6. Differences between 3 methods, HAMM1, HAMM2 and HAMM3, are that each uses difference optimization method to find those parameters. As Jelali's HAMM1 uses grid search, Lee's HAMM2 uses multi-start adaptive random search and Karra's HAMM3 uses an Extended ARMAX model.

$$S = \frac{f_s + f_d}{2} \tag{5}$$

$$J = \frac{f_s - f_d}{2} \tag{6}$$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

For the performance of each detection method, in 2010, Jelali collaborated with Choudhury, He, Thornhill and Huang provided a benchmark called, international stiction data base (ISDB), consisted of 93 different data sets from different process industries, including chemicals, pulp and paper mills, commercial building, and metal processing. Examples of the database are shown in Figure 9. And in 2020 Kamaruddin et al. [21] filtered some of those data sets and provided a comparision of those techniques as shown in Table 1.

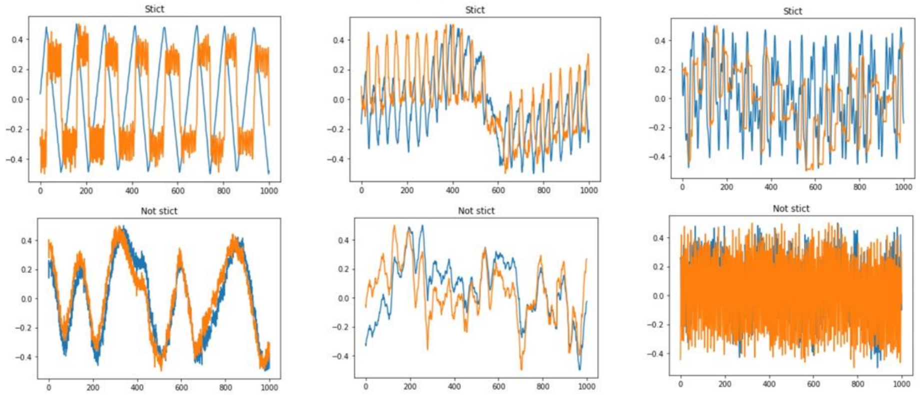


Figure 9. Examples of the data in ISDB

As shown in Table 1, most methods provide rather unsatisfactory performance with F1 score only around 0.5 – 0.7. And while BIC method gives much more acceptable result, as discussed earlier, it requires at least 1,024 data points without abrupt change which many plants cannot provide [22].

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Table 1. Comparison of classical detection methods

Method	F1 score
BIC	0.84
CORR	0.51
HIST	0.58
RELAY	0.63
CURVE	0.58
AREA	0.60
HAMM2	0.64
HAMM3	0.66

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.3 Machine learning detection methods

As the main problem of traditional methods is that, even though characteristics of stiction signal are known as discussed in previous sections, real world data are usually corrupted by noise, disturbances and controller commands to the point that detecting those characteristics is impractical in many cases. As of today, there is also no universal way to clean the data without losing some of the stiction features [23]. On the other hand, it is almost effortless to simulate corrupted real-world-like data. So, in an expectation that with enough corrupted data, with and without stiction traits, a well-built machine learning model should be able to learn and extract stiction characteristics and classify the data, recent trend in valve stiction detection is shifted to machine learning approaches.

One of the first machine learning detection methods was proposed about a decade ago. In 2009, Farenzena [24] purposed an idea of estimate valve stiction parameters based on pattern recognition. As aforementioned the main challenge is that there is practically no real-world dataset to train a model. So, he, as well as every method that will be mentioned afterward in this section, simulated their own training and test set with Choudhury's stiction model by varying controller and process parameters, stiction parameters: S and J, as well as variance of white noise. More details of the data generation process are shown in Appendix A. The model is a basic regression artificial neural network model (ANN) with 1 10-neurons hidden layer with difference between the maximum and minimum value in PV and OP as well as number of zero-crossings in the autocovariance function as inputs. And while he concluded that the model performed slightly worse than traditional methods, it showed some potential especially when considering its simplicity and low computational time.

In the same year, Zabiri et al. proposed 2 research on the subject [25] [26]. In those papers, they compare performance on estimating the stiction parameters of different types of neural network including basic feedforward backpropagation, nonlinear autoregressive network with exogenous inputs (NARX) and recurrent networks with MV and PV as inputs. In their conclusion, only recurrent networks were able to give acceptable performance for the simulation data, but they also noted that it was highly unlikely to be able to perform with real-world data due to low robustness in the presence of external disturbances.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Nevertheless, as in real process MV data is commonly unknown, in 2012, Venceslau et al. [27] proposed a model using OP and PV data as inputs. In their model 40-sized seqeunce of d-value defined as eq. 7 is used as the model input and the model is an ANN with 1 15-neurons hidden layer with estimation of S and J as outputs. In 2019, Amiruddin et al. [28] adopt the idea and developed their own ANN model (SDN) with 2 20-neural hidden layers and 500-sized d-value as input. However, in this case, insted of estimate value of S and J it is a classification model that return whether a process has valve stiction problem or not.

$$D_i = \sqrt{(x_{op,i} - \frac{1}{N}\sum_{i=1}^{n} x_{op,i})^2 + (x_{pv,i} - \frac{1}{N}\sum_{i=1}^{n} x_{pv,i})^2} \tag{7}$$

Then In 2019, Dambos et al. [29] propose a new method to prepare an input for ANN models. Inspired by Choudhury PV(OP) plot method, instead of using d-value as input, they used an 8x8 pixels of PV(OP) plot from 1,000 data points each as shown in Figure 10. Although there is not extensive tested, they believed that this method should be more robust to signals with different lengths and sampling time, since they do not change the shape of PV(OP) plot. In 2020 Kamaruddin et al. [21]. proposed another way to process picture-based inputs, butterfly shape-based detection (BSD). The main idea of this method is similar to Dambos's as they did not use the fix size time series to avoid the signal length problem. But, instead of PV(OP) plot, they proposed a relationship between |OP(k-1)-PV(k-1)|(PV(k)) which theoretically should form a butterfly shape in cases of stiction as shown in Figure 11. Then after they confirmed the shape using Identification of Round Objects Method (IROM) method, they use the picture as an input for a convolutional neural network (CNN) to classify degree of stiction (weak, moderate, or strong). The structure of their model is shown as in Table 2.

จุฬาลงกรณ์มหาวิทยาลัย

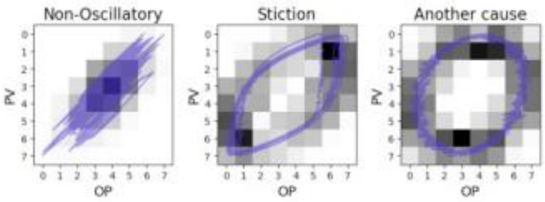


Figure 10. 8x8 pixels input for Dambos model

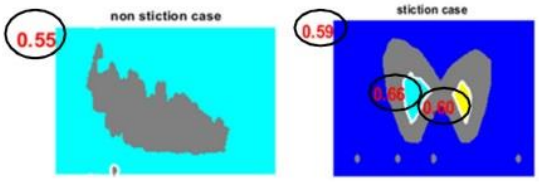


Figure 11. Input shape of BSD method

Table 2. Structure of Kamaruddin BSN CNN model

Layer Name (Number)	Filter Size	Number of Filters	Layer Name (Number)	Size	Movement Method	Movement Size
Convo (1)	5	20	Max pooling (1)	2	Stride	2
Convo (2)	5	40	Max pooling (2)	2	Stride	2
Convo (3)	5	80	Max pooling (3)	2	Stride	2

Lastly in 2021, Zhang et al. [30] proposed a way to utilize a more complex model for valve stiction detection. They combined 2 CNN models with structure as show in Table 3, first model (Net_{fix}) used Euclidean distance of PV and OP with fixed range, size 50x50 as an equation shown in eq.8. The second one ($\text{Net}_{\text{unfix}}$) used PV(OP) plot like in Dambos's work but with 32 x 32 pixels.

Table 3. Structure of Zhang CNN models

Model	Main Setting
Net_{fix}	Net: Cov2d (3,8,5,1) > ReLu() > Maxpool2d (2,2) > Conv2d (8,25,5,1) > ReLU() > Maxpool2d (2,2) > Fc (625,120) > ReLU()> Fc(120,84)>ReLU()>FC(82,2) Data: Format: 2D- matrix Input Size: 50 x 50 Training Timescale: 200,300,400,600 Test Timescale: 50, 75, 100, 200

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$\text{Net}_{\text{unfix}}$	Net: Cov2d (3,8,5,1) > ReLu() > Maxpool2d (2,2) > Conv2d (8,25,5,1) > ReLU() > Maxpool2d (2,2) > Fc (625,120) > ReLU()> Fc(120,84)>ReLU()>FC(82,2) Data: Format: .jpg Image Size: 32 x 32 pixels Training Timescale: 200,300,400,600 Test Timescale: 50, 75, 100, 200

$$m_{uv}^{i} = \sqrt{\frac{\sum_{j=1}^{|s_i|/n} (x_{op}^{u \leftarrow j} - x_{pv}^{v \leftarrow j})^2}{|s_i| / n}} \tag{8}$$

In this research, Zhang compared the results of the combined model, each individual model, result using other machine learning model like SVM and Lenet-5 as well as Amiruddin’s SDN and Kamaruddin’s BSD. Then, they claimed that their model had the best performance as shown in Table 4. In addition of the claim there are 2 more interesting points from the table, firstly while method like Randomforest (RF) and Xgboost failed to give a result, BSD, SDN and using LENet-5 model structure all provide F1 score higher than 0.7 without many limitations like BIC method. Secondly, when comparing 3 Zhang’s models, the model with unfix input give F1 score closed to the final model and much better than the model with fix input.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 4. Comparison of machine learning detection methods

Method	Stiction			Non-Stiction		
	P	R	F1	P	R	F1
Zhang	0.87	0.9	0.89	0.9	0.87	0.88
LR	0.57	0.53	0.55	0.56	0.6	0.58
RF	0.74	0.47	0.57	0.61	0.83	0.7
SVM	0.68	0.7	0.69	0.69	0.67	0.68
Xgboost	0.63	0.4	0.49	0.56	0.77	0.65
LeNet-5	0.8	0.8	0.8	0.8	0.8	0.8
BSN	0.75	0.73	0.74	0.77	0.79	0.78
SDN	0.8	0.8	0.8	0.73	0.73	0.73
Net_{fix}	0.76	0.63	0.69	0.69	0.8	0.74
$\text{Net}_{\text{unfix}}$	0.86	0.77	0.85	0.81	0.97	0.88

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER IV

METHODOLOGY

This chapter focuses on our stiction detection approach. We begin by presenting the data generation method utilized to train our deep learning model. Subsequently, we delve into the data preprocessing step, which aims to improve interpretability and minimize reliance on simulating the training set. Following that, we showcase the architecture of our Convolutional Neural Network model. Finally, we discuss the test set and evaluation criteria employed in our study.

4.1 Data generation for model training

To train our CNN model, we employed simulated data, following standard practices in the field. Our simulation system, illustrated in Fig. 1, consisted of a first-order process with predefined gain and time constant values of 1 and 0.2 minutes, respectively. Choudhury's valve model represented the valve transfer function, while a simplified form of the PI controller, as described in eq.9, represented the controller. The parameter values used in the simulations are presented in Table 5. Specifically, for the stiction case, we varied the controller gain between 2 and 5 with a step size of 1, the integral gain between 300 and 500 with a step size of 100, and the white noise variance between 0 and 0.05 with a step size of 0.005. These parameter settings aimed to simulate well-tuned and slightly aggressive systems with moderate white noise. Additionally, we varied the stiction parameters fs (between 2 and 5) and f_d (between 1 and f_s) with a step size of 1 for both.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

For the stiction case, we extended the range of the first three parameters to 2-7, 200-700, and 0-0.1 for the controller gain, integral gain, and white noise variance, respectively. These adjustments allowed us to simulate more aggressively tuned systems with increased levels of noise. To ensure a stable initial state, we initialized the process setpoint (SP) to 0 and transitioned it to 1 at the beginning of each simulation. Data collection commenced from the 1st minute of the simulation, to ensure a stable initial state, and involved recording the output of the process variable (PV) and the controller output (OP) at a sampling rate of 0.01 minutes. Consequently, each simulation generated time series data for both PV and OP, with a length of 800 data points for each variable.

Importantly, we note that while specific parameter choices were made in our simulations, these choices should not significantly impact the model's performance as long as they are reasonably selected. This allows for flexibility in adapting the simulation settings to match real-world scenarios and diverse operating conditions.

$$OP(S) = K_p E(S) + \frac{K_i E(S)}{S} \tag{9}$$

Table 5. Parameters of data generation

K_c	K_i	f_s	f_d	$\text{Variance}_{\text{noise}}$
[2:1:7]	[200:100:600]	0	0	[0:0.01:0.1]
[2:1:5]	[300:100:500]	[2:1:5]	[1:1:f_s]	[0:005:0.05]

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2 Data preprocessing

For the data preprocessing step, we employed a multi-step approach. Firstly, we estimated the oscillation frequency of the data using the PSD on PV. The primary idea is that ACF can preserve oscillation characteristics of the data while provide a better representation of the signal-to-noise ratio compared to raw data itself and Fourier transform was used to decompose the signal into frequencies, making it easier to distinguish and measure oscillatory components. We then applied a high-pass filter to clean out low-frequency components (less than 0.8 of the oscillation frequency we found) of the data. If no oscillation frequency was found, we skipped this step.

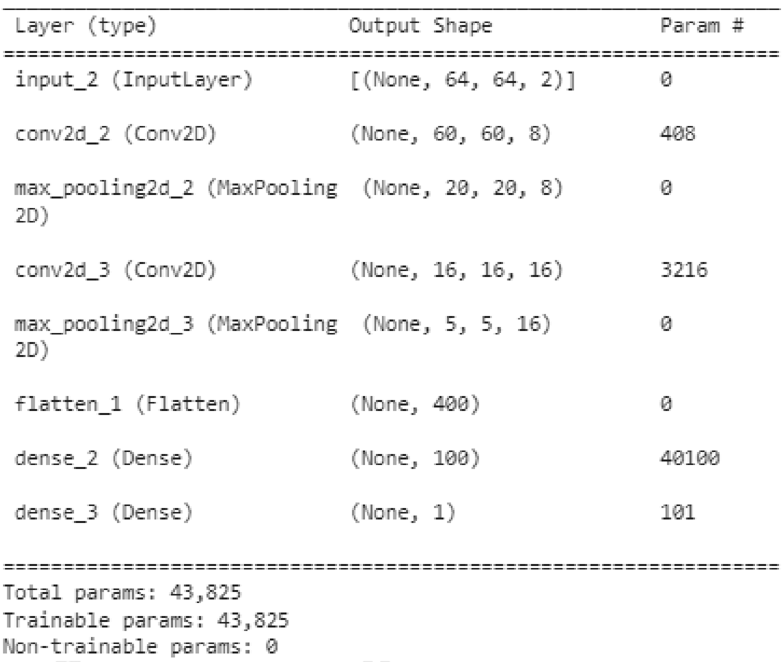
Next, we use wavelet reconstruction to de-noise the filtered data. The idea is we decompose the data to coefficient array of wavelet in different scales. Then we identify the scales with insignificant coefficients by setting a threshold for the standard deviation of each coefficient array. If the standard deviation is less than the threshold, we set the entire array to zeros. Otherwise, we keep the original array. Lastly reconstruct the signal. This step can de-noise our data while preserving striction characteristic due to Lipschitz regularity theory as mentioned in the previous section. For this work, we use Daubechies 2 as wavelet basis, do 7 levels of decomposition and use the standard deviation threshold = 0.5.

For the last step of the Data preprocessing, we create grayscale PV(OP) plots of both filtered data before the reconstruction and reconstruction data. Then we convert each of them to a 64x64 array. So, for each data series we get an input for our CNN model with a shape (64,64,2) regardless of data series length. And for the training of the model, the simulated data was with split 80:20 ratio for train and validation.

4.3 Convolutional neural network model architecture

For this study, we implemented a CNN to classify whether a valve is experiencing stiction or not, based on the preprocessed data generated as described above. It consists of two convolutional layers, each using a 5x5 kernel size and having 16 and 32 output filters, respectively. These convolutional layers are followed by max-pooling layers with a 3x3 pool size, which help downsample the feature maps and extract the most relevant information. Subsequently, we utilize two fully connected layers with 100 and 1 neurons, respectively, with a dropout layer in between to mitigate overfitting by randomly disabling some neurons during training.

To introduce non-linearity and enable effective learning, Rectified Linear Unit (ReLU) activation functions are employed for all layers except the output layer. The output layer uses a sigmoid activation function to produce a binary classification output indicating the presence or absence of stiction. The CNN model is trained using the Adam optimizer, which is known for its efficiency in handling large datasets. We utilize a binary cross-entropy loss function to measure the difference between the predicted and actual labels. The model is trained for 100 epochs with a batch size of 64, allowing it to iteratively adjust its parameters to minimize the loss and improve classification performance.



```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_2 (InputLayer)        [(None, 64, 64, 2)]       0

 conv2d_2 (Conv2D)           (None, 60, 60, 8)         408

 max_pooling2d_2 (MaxPooling  (None, 20, 20, 8)        0
 2D)

 conv2d_3 (Conv2D)           (None, 16, 16, 16)        3216

 max_pooling2d_3 (MaxPooling  (None, 5, 5, 16)         0
 2D)

 flatten_1 (Flatten)         (None, 400)               0

 dense_2 (Dense)             (None, 100)               40100

 dense_3 (Dense)             (None, 1)                 101

=================================================================
Total params: 43,825
Trainable params: 43,825
Non-trainable params: 0
```

CHULALONGKORN UNIVERSITY

Figure 12. Convolutional neural network model architecture

4.4 Test set and evaluation

To evaluate our model, we used 29 data from the ISDB dataset which have been manually labelled and were used to compare traditional features based in Mohieddine’s study. The test set consists of data from 4 industries, several types of control loops including flow, temperature, level, pressure, concentration, and gauge and with varied sampling time. Of the 29 data, 15 loops have the stiction problem, whereas the other 14 have other or no problem.

For evaluation, our primary evaluation metric to assess the performance of our CNN model is F1 score which is a commonly used evaluation metric in machine learning and classification tasks. It is a measure of a model's performance that combines both precision and recall into a single value. The F1 score ranges from 0 to 1, with a value of 1 indicating perfect precision and recall, and a value of 0 indicating the opposite. By considering both precision and recall, the F1 score provides a balanced evaluation of a model's performance on both classes and is often preferred over other metrics such as accuracy. The score is calculated as in eq. 10, eq. 11 and eq. 12 where TP, FP and FN are a number of true positive, false positive and false negative cases, respectively.

$$Precision = TP/(TP + FP) \tag{10}$$

$$Recall = TP/(TP + FN) \tag{11}$$

$$\mathrm{F1} = 2\ (Precision * Recall)/(Precision + Recall) \tag{12}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 6. Test dataset

Loop Name	Comments	Loop Name	Comments
CHEM 1	Striction	CHEM 32	Striction
CHEM 2	Stiction	CHEM 33	Disturbance
CHEM 3	Quantisation	CHEM 40	No clear oscillation
CHEM 6	Stiction	CHEM 54	No clear oscillation
CHEM 10	Stiction	CHEM 62	No clear oscillation
CHEM 11	Stiction	PAP 2	Stiction
CHEM 12	Stiction	PAP 4	Deadzone and tight tuning
CHEM 13	Faulty sensor	PAP 5	Stiction
CHEM 14	Faulty sensor	PAP 7	Disturbance
CHEM 16	Interaction	PAP 9	No striction
CHEM 18	Stiction	MIN1	Stiction
CHEM 23	Stiction	MET1	Disturbance
CHEM 24	Stiction	MET2	Disturbance
CHEM 28	Stiction	MET3	No oscillation
CHEM 29	Stiction		

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER V

RESULT AND DISCUSSION

In this chapter, the results and discussions are structured into six sections for a comprehensive analysis. Firstly, we delve into the outcome of the data preprocessing step and its impact on interpretability. Secondly, we explore the overall performance of the proposed approach, compare it with other methods, and look at their fault cases. Next, we conduct a sensitivity analysis to assess the robustness of our method. Lastly, we address the exploration of alternative deep learning models.

5.1 Model input interpretability

To gain insights into the data preprocessing technique, we performed an analysis of the PV(OP) plots obtained from both non-strict and strict cases. Figure 13 illustrates the results of this analysis, where we compared the raw data with the plots after applying the PSD and high-pass filter. Initially, when examining the raw data, the presence of non-stationary slowly varying trends made it challenging to discern the plot's shape accurately. However, by applying the PSD and high-pass filter, we successfully reduced these effects, resulting in a much clearer and more comprehensible plot. Surprisingly, we observed that there was little discernible difference between the shapes of the PV(OP) plots in the non-strict and strict cases, posing a significant challenge for accurate classification.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

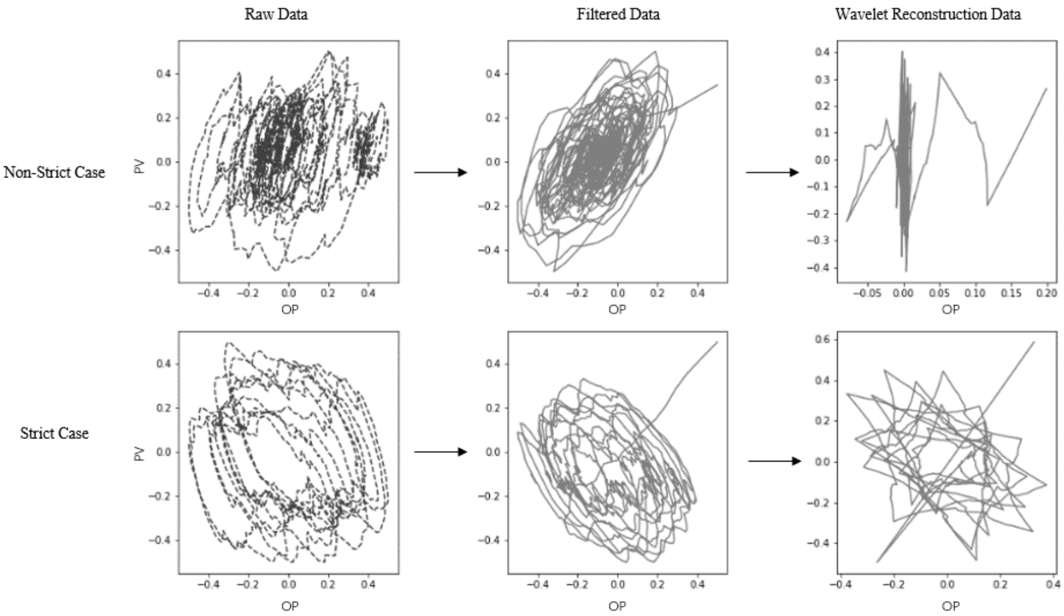


Figure 13. PV(OP) Plots obtained from both non-strict and strict cases

To overcome this challenge, we employed the wavelet reconstruction data, which provided enhanced differentiation between the two cases. In the non-strict case, the denoising effect of the wavelet reconstruction caused the PV(OP) plot to lose its shape, making it distinct from the strict case. We validated this observation across both simulated and real data in the test set, reinforcing the robustness of our approach to variations in the training set generation. Furthermore, we discovered that setting the threshold for wavelet reconstruction to 0.5 resulted in the clearest distinction between the two cases and yielded the best overall performance. This finding elucidates the reason behind its superior performance and highlights the significance of utilizing both the filtered and wavelet reconstruction PV(OP) plots as inputs to the model. The contrasting shapes of these plots provide crucial insights that improve the model's ability to accurately classify stiction behavior.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

5.2 Performance evaluation

The performance evaluation of our proposed model on the test set yielded compelling results, showcasing its effectiveness in stiction detection. The model achieved an impressive F1 score of 0.90, indicating a harmonious balance between precision and recall. With a precision of 0.87 and a recall of 0.93, the model exhibited a high level of accuracy in correctly identifying instances of stiction. Notably, the model exhibited minimal errors, with only two false positives and one false negative. The Confusion Matrix in Figure 14 provides a detailed breakdown of the model's classification results, showcasing the true positives, true negatives, false positives, and false negatives. This visual representation further emphasizes the model's accurate performance in detecting stiction.

To assess the model's performance in greater detail, we computed an accuracy rate of 90%, highlighting its remarkable capability to correctly classify the majority of instances. This performance is further demonstrated through a direct comparison with the widely utilized BIC metho. The BIC method attained an F1 score of 0.80 and an accuracy rate of 76% on the same dataset. It is important to note that the BIC method yielded six false positives and one false negative, indicating a less precise identification of stiction occurrences compared to our proposed model.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

		Predicted	
		Negative	Positive
Actual	Negative	12	2
	Positive	1	14

Figure 14. Confusion matrix

Another support to the idea can be seen as we investigated feature map visualization of the model. As shown in examples of the outputs of second convolutional layers of the model in Figure 15-19, as it demonstrates that in addition to the ellipse shape of the plot, model also looking for the difference between wavelet reconstruction (red) and high pass filtered data (blue) as clearly shown in filter number three.

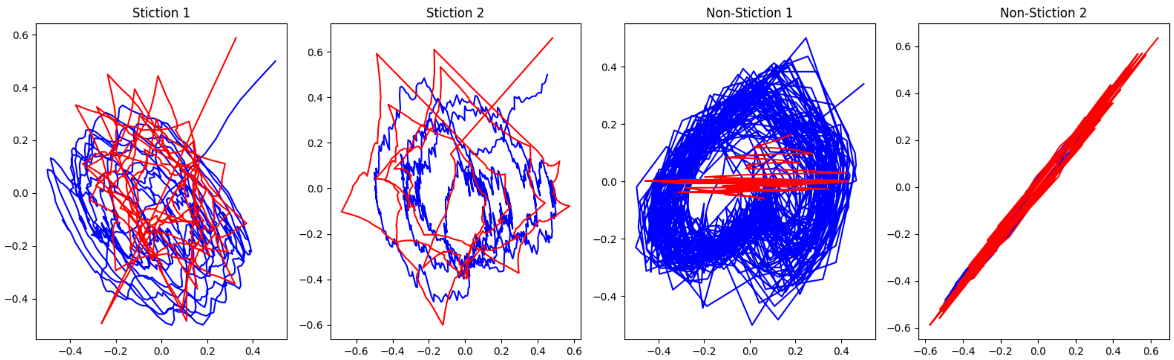


Figure 15. Model input examples for feature map visualization

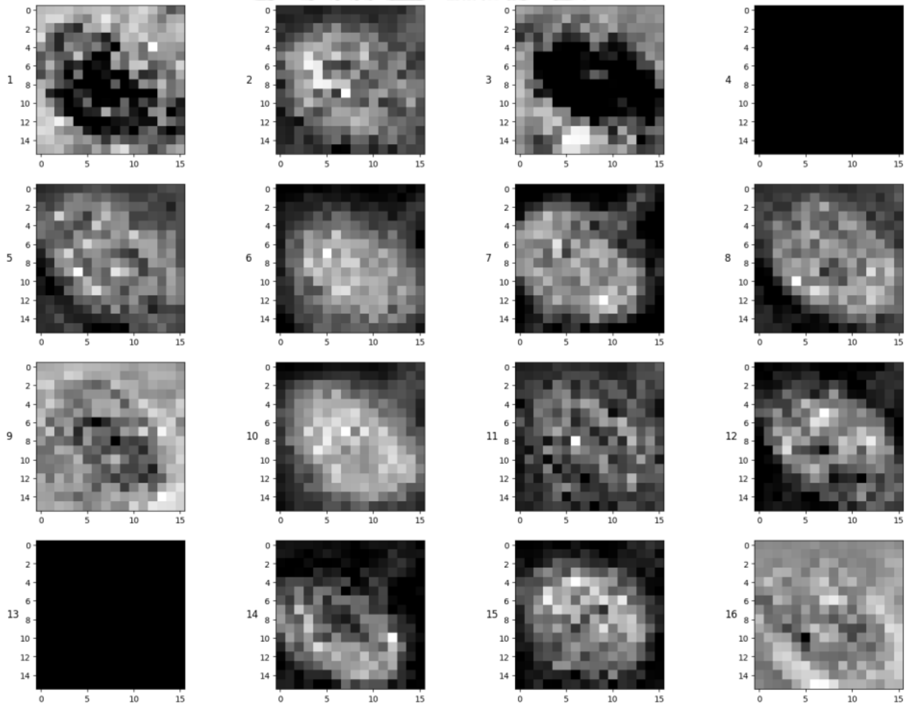


Figure 16. Second convolutional layers output of Stiction 1 example

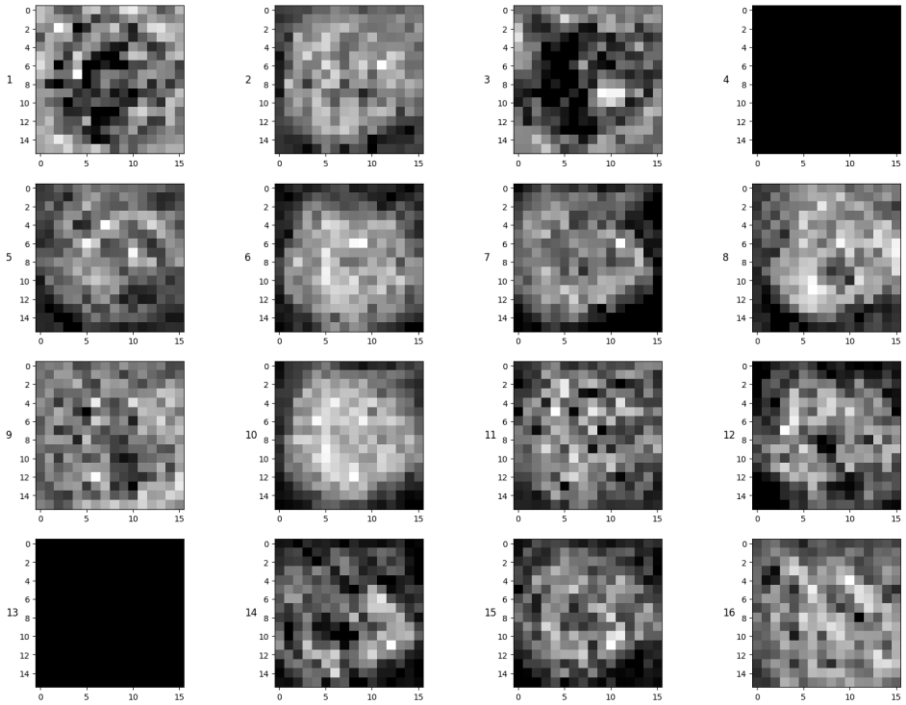


Figure 17. Second convolutional layers output of Stiction 2 example

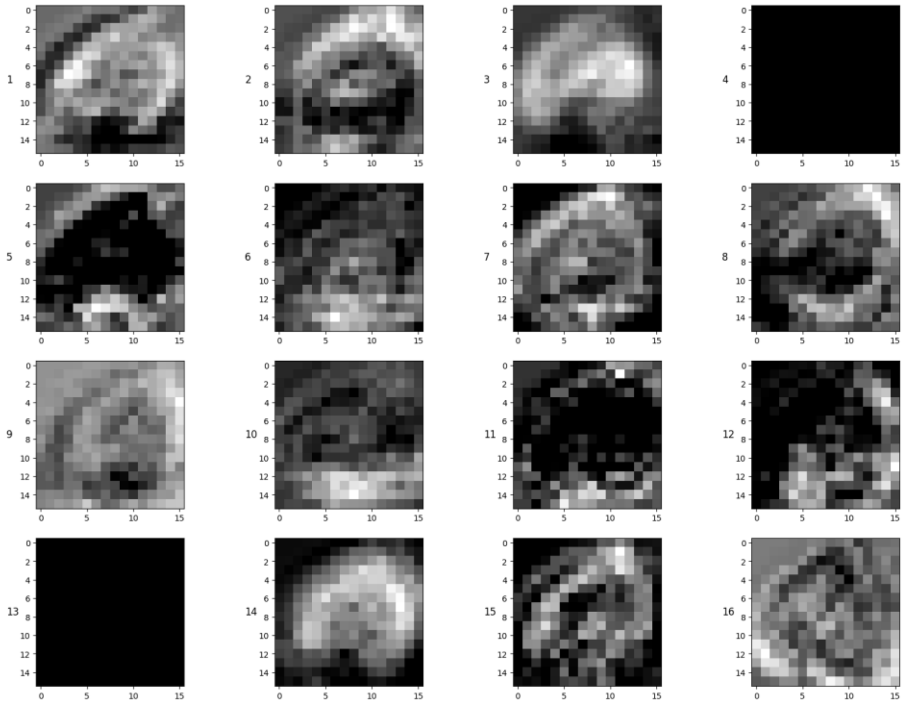


Figure 18. Second convolutional layers output of Non-Stiction 1 example

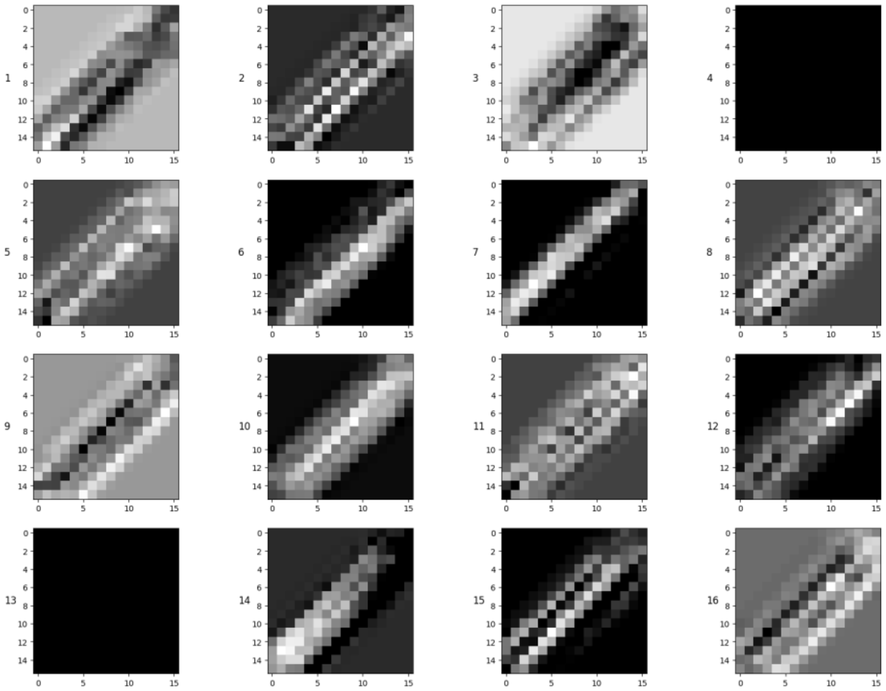


Figure 19. Second convolutional layers output of Non-Stiction 2 example

Furthermore, the Receiver Operating Characteristic (ROC) curve, illustrated in Figure 20, highlights the trade-off between the true positive rate and the false positive rate across different classification thresholds. The curve exhibits a high Area Under the Curve (AUC) of 0.92, indicating the model's excellent ability to discriminate between instances of valve stiction and non-stiction.

The AUC value of 0.92 demonstrates that the proposed method achieves a high true positive rate while maintaining a low false positive rate. This suggests that the model has a strong discriminatory power, effectively distinguishing between cases of valve stiction and non-stiction. The ROC curve, located in the upper-left quadrant, further supports the high performance of the model, as it approaches the ideal point of perfect classification.

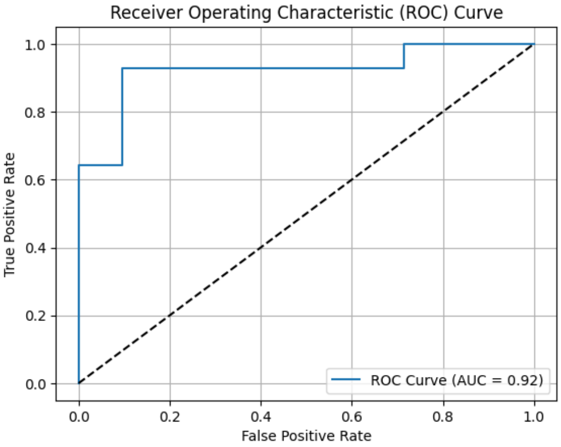


Figure 20. ROC curve

These results substantiate the superiority of our proposed method over other approaches in stiction detection. The model's higher F1 score, accuracy rate, and minimized false positives and false negatives demonstrate its potential to outperform existing methods, thereby offering a promising solution to address the challenges associated with stiction detection in process systems.

5.3 Result comparison

In this section we compared results of our proposed method with some existing ideas from Zhang's research albeit using our generated training set and using entire data length for both training and test model instead of only the first fixed length. Specifically, we compared three different types of model input: our proposed approach utilizing two layers of grayscale modified PV(OP) plots, the PV(OP) plots of raw data as used in Zhang's research, and the PV(OP) plots of high pass filtered data. For the classifier, we employed SVM, RF, and CNN. It is worth noting that for the other two input types, we utilized the LeNet-5 architecture, which was consistent with Zhang's research.

For other models, we utilized the scikit-learn library to implement both the SVM model and the RF Classifier. For the SVM model, the sklearn.svm.SVC class was employed with default settings. It utilizes the Radial Basis Function (RBF) kernel to

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

capture non-linear relationships in the data. The regularization parameter (C) was set to its default value of 1.0, while the gamma parameter was set to 'scale' to automatically calculate its value based on the input data. Regarding the Random Forest Classifier, we used the sklearn.ensemble. RandomForestClassifier class with default settings as well. This classifier constructs an ensemble of 100 decision trees, where each tree is built using a random subset of features and training data to ensure diversity. The random_state parameter was set to 42 to seed the random number generator and ensure result reproducibility. Default values for other settings were also maintained.

Based on the comparison of F1 scores presented in Table 7, it is evident that our proposed modified PV(OP) input approach, in conjunction with the CNN classifier, achieved the highest F1 score of 0.9. While other methods demonstrated similar performance levels, the combination of our modified approach and the CNN classifier proved to be the most effective in accurately detecting valve stiction. This finding highlights the synergistic relationship between our modified input approach and the CNN classifier, resulting in superior performance for valve stiction detection.

Table 7. Comparision of F1 Score with different models and inputs

Method	Our proposed input	High pass filtered input	Raw data input
CNN	0.90	0.77	0.79
SVM	0.77	0.71	0.78
RF	0.73	0.74	0.71

5.4 Fault cases analysis

To gain deeper insights into the classification process, we conducted an analysis of the fault cases classified by each input type using the CNN classifier and the traditional BIC method. Table 8 provides an overview of all the fault cases, with

PV(OP) plots of raw data, high pass filtered data, and wavelet reconstruction data represented by green, blue, and red lines, respectively.

Starting with the cases where our proposed method outperformed the others, such as CHEM14, CHEM16 and CHEM54, these demonstrate the desired effects of our modified input approach. Without the wavelet reconstruction data, these plots exhibited clear indications of stiction, resembling textbook stiction cases. Similarly, in the case of PAP5, although the shape was ambiguous, the preservation of the shape through wavelet reconstruction enabled correct classification.

On the other hand, let's examine the cases where our model misclassified the faults. In the cases of CHEM13 and CHEM29, our approach exhibited weaknesses as other methods had no difficulty in classifying them correctly. Particularly in the case of CHEM29, the absence of wavelet reconstruction data resulted in an unmistakable stiction plot, but the reconstruction process filtered out the shape, leading to misclassification. In future research with larger datasets, it would be worth reevaluating the hyperparameters for the wavelet reconstruction threshold to address such challenges.

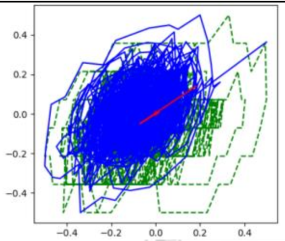
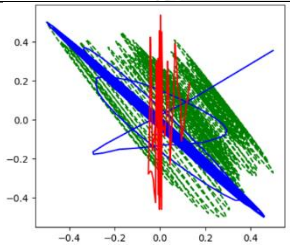
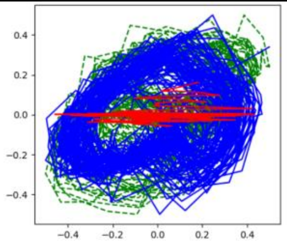
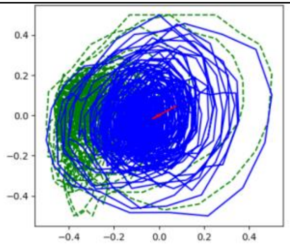
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

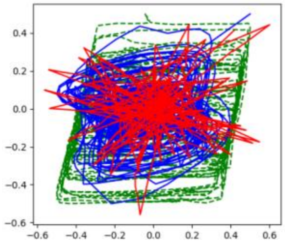
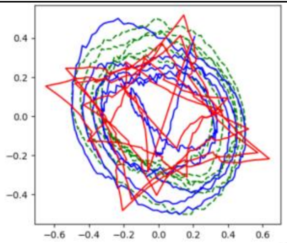
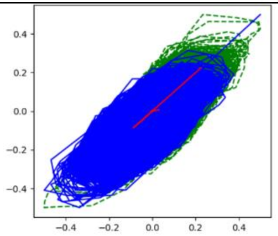
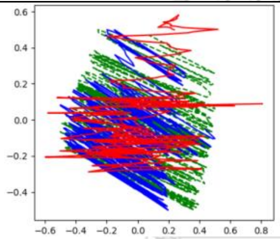
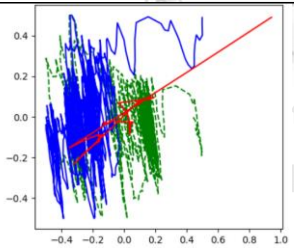
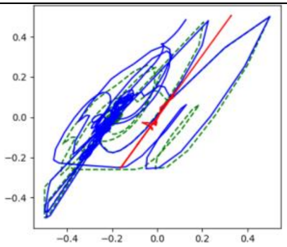
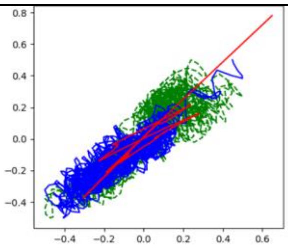
Furthermore, interesting findings emerged when comparing the results between the other input types. Although the model with raw data input slightly outperformed the one with high pass filtered input, the actions of the former were more difficult to comprehend. For example, in the case of CHEM3, a well-known challenging case for classification where most traditional methods, except BIC, tend to misclassify, the model's lack of exposure to similar shapes during training could lead to accidental correct predictions, especially considering its incorrect prediction of non-stiction for CHEM24 with the similar raw data PV(OP) shape. This underscores the notion that simulating data to accurately represent the vast array of real-world scenarios is nearly impossible, particularly without appropriate data preprocessing.

Additionally, there were cases where only the BIC method yielded incorrect predictions, likely due to inaccuracies in the Gaussian and linearity tests, which our approach successfully avoided.

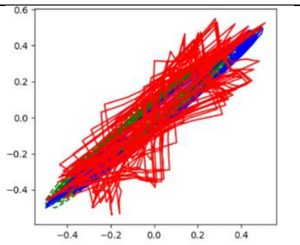
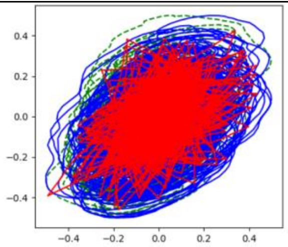
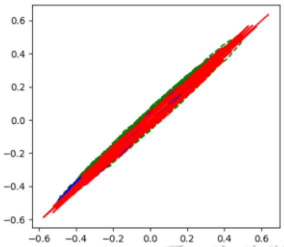
Overall, this fault case analysis provides valuable insights into the strengths and weaknesses of our proposed method and the different input types. It underscores the importance of considering the interpretability of the model's predictions and the challenges associated with simulating real-world data.

Table 8. Fault cases of different methods

Data	PV(OP) plot	Striction	Proposed input	High pass filtered input	Raw data input	BIC
CHEM3		NO	NO	YES	NO	NO
CHEM13		NO	YES	NO	NO	NO
CHEM14		NO	NO	YES	NO	YES
CHEM16		NO	NO	NO	YES	YES

CHEM24		YES	YES	NO	NO	YES
CHEM28		YES	YES	YES	YES	NO
CHEM29		YES	NO	YES	YES	YES
CHEM33		NO	NO	YES	NO	YES
CHEM40		NO	NO	YES	YES	NO
CHEM54		NO	NO	YES	YES	YES
CHEM62		NO	NO	YES	YES	NO

ณ์มหาวิทยาลัย
KORN UNIVERSITY

PAP4		NO	YES	NO	YES	YES
PAP5		YES	YES	NO	NO	YES
PAP9		NO	NO	NO	NO	YES

5.5 Sensitivity analysis

For sensitivity analysis, we explored the influence of some model parameters on its performance. We varied various parameters including the number of training epochs within a reasonable range of 50 to 200, number of filters in each CNN layer and different values of the learning rate, ranging from 0.001 to 0.1. As shown in Figure 21-23, except for the too high leaning rate and too few filters, we observed the adjusting the numbers had a negligible impact on the model's accuracy. The model consistently achieved an average F1 score of around the same across different model parameters. These findings suggest that our model is robust and stable across a wide range of parameter values, further supporting its reliability for stiction detection tasks.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

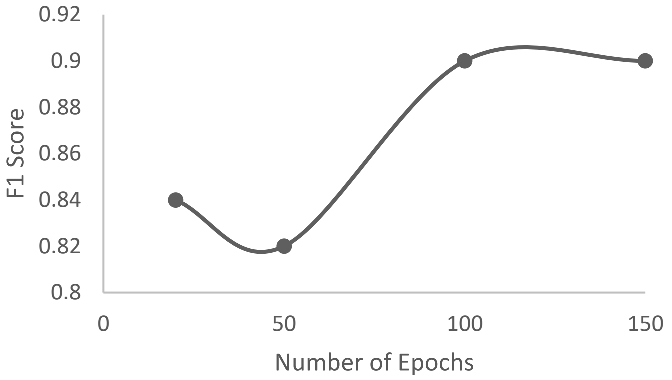


Figure 21. Model performance with different number of epochs

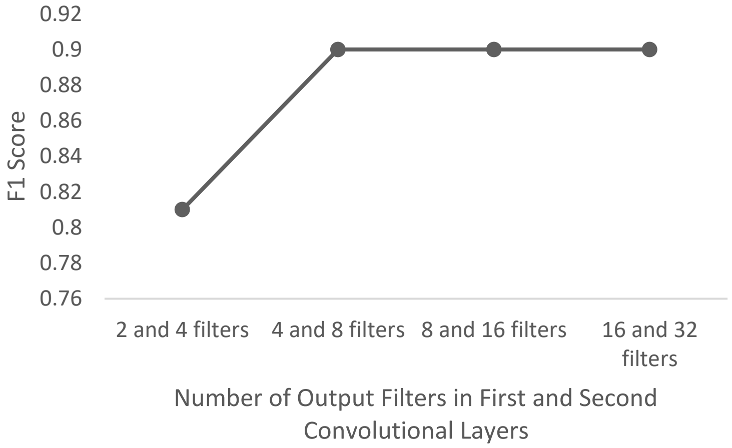


Figure 22. Model performance with different number of output filters in convolutional layers

CHULALONGKORN UNIVERSITY

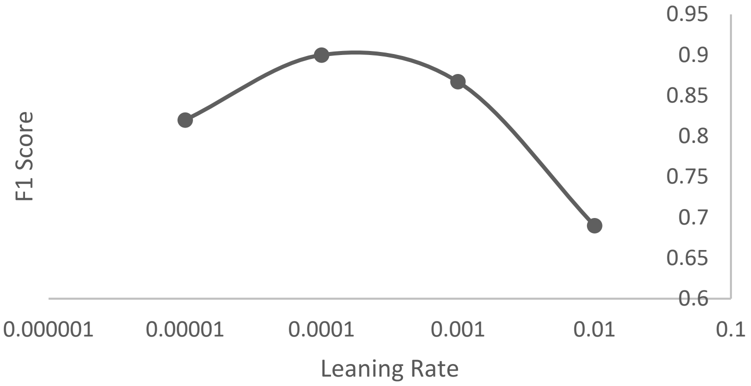


Figure 23. Model performance with different leaning rate

While we discovered that various factors had little impact on the performance of our model. However, one intriguing factor that significantly influenced the results was the threshold for wavelet reconstruction, as depicted in Figure 10. By varying the threshold between 0.3 and 0.7, we observed changes in performance within the range of 0.76 to 0.9. Notably, a threshold of 0.5 yielded the best result. The threshold serves as a critical parameter in determining which wavelet resolutions should be filtered out. Setting the threshold too high may inadvertently allow unwanted factors like white noise to pass through, while setting it too low can lead to the removal of crucial valve stiction characteristics. It is important to note that although a threshold of 0.5 provided the optimal result in our study, it may not necessarily be the most suitable setting for real-world scenarios. Further data collection and analysis would be necessary to determine the appropriate threshold for such scenarios.

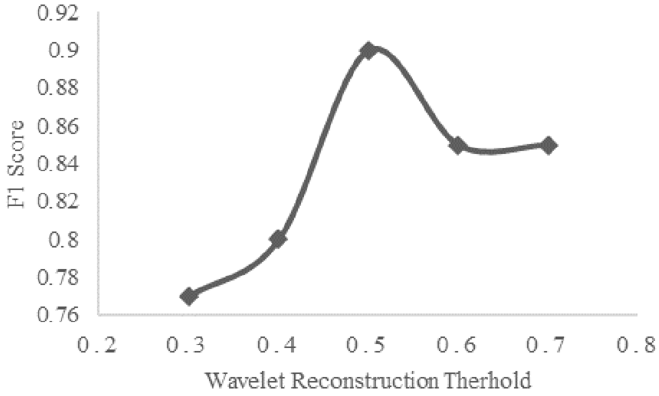


Figure 24. Model performance with different wavelet reconstruction threshold

Another aspect we examine in this section is the impact of randomness in the data generation process. While we aimed to create a training set that best represents the target concept, inherent randomness, especially from the generated noise, introduced variations in the quality of the training set and consequently affected the model's performance. As shown in Figure 21, out of the 20 simulation sets, the model did not consistently achieve the outstanding F1 score of 0.9; however, it still demonstrated relatively strong performance with an average F1 score of 0.855 and a standard deviation (SD) of 0.03.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

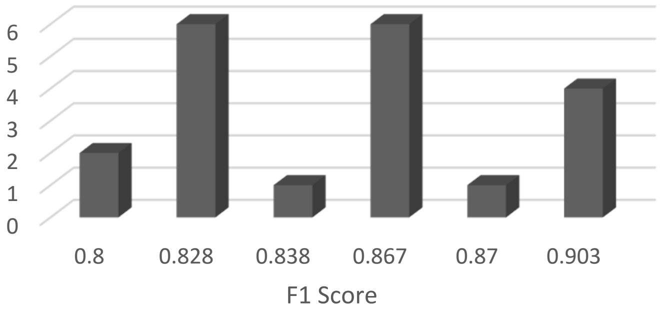


Figure 25. Model performance distribution with 20 different training sets

Furthermore, we conducted experiments by varying the process time constant value from 0.2 minutes to 0.1 and 0.4 minutes, each with 10 simulation sets. The average F1 scores for these cases were 0.825 and 0.842, respectively, with SD values of 0.002 for both cases. The slight decrease in performance can be attributed to the specific parameter selections, such as the range of the controller gain, which were tailored for the original time constant value.

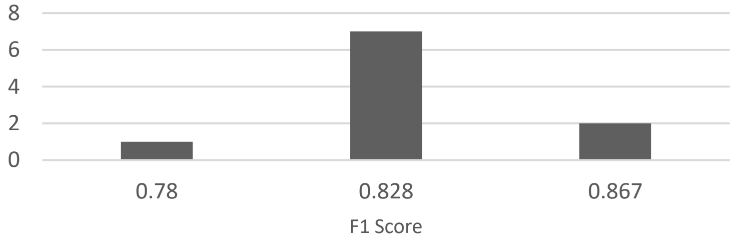


Figure 26. Model performance distribution with process time constant = 0.1 minutes for10 different training sets

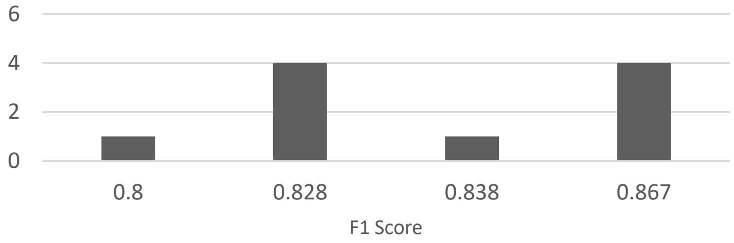


Figure 27. Model performance distribution with process time constant = 0.4 minutes for10 different training sets

5.6 Other neural networks

In our exploration, we investigated the performance of alternative models to compare them against our proposed model. Specifically, we considered a simple fully connected artificial neural network (ANN) and a recurrent neural network (RNN) model with long short-term memory (LSTM) cells.

However, when evaluating these alternative models, we found that their performance was relatively lower compared to our proposed model. The fully connected ANN achieved F1 scores ranging from 0.6 to 0.7, indicating limitations in capturing the complex patterns and relationships inherent in the valve stiction detection task. This outcome emphasized the significance of leveraging the specialized architecture and capabilities offered by convolutional neural networks (CNNs) for addressing this specific problem.

Regarding the RNN model with LSTM cells, which is specifically designed for handling sequential time series data, we encountered significant challenges when adapting it to the valve stiction detection task. One of the main difficulties we faced was determining the appropriate sequence length to feed into the model. Given the considerable variations in data length, oscillation frequency, and sampling time between the test set and real-world scenarios, finding a uniform sequence length that adequately represented the data proved to be problematic. As a result, the performance of the RNN model was hindered, yielding F1 scores similar to those obtained with the fully connected ANN.

To address the challenges associated with varying data lengths, we attempted a different approach. Instead of using a fixed time window, such as 200 samples from 200-time steps, which could be too short or too long for individual data and might not capture the characteristic oscillation patterns effectively, we introduced flexibility in the sampling window. By leveraging the information obtained from the power

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

spectral density (PSD) analysis to determine the data's oscillation frequency, we calculated a suitable window length, approximately 3-4 times the period of the oscillation. We then interpolated the collected data to form a 200-sample length input, which we fed into our model. Although this approach did not yield results comparable to our CNN model, we believe it holds potential and could be worth further investigation in future research endeavors.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER VI

CONCLUSION

In this study, we have proposed a new method for detecting valve stiction in control loops based on a combination of wavelet reconstruction and convolutional neural networks. Our proposed method showed strong performance in detecting stiction in various types of control loops and achieved an F1 score of 0.90 on a test set of 29 data from the ISDB dataset. Compared to the traditional BIC method, our proposed method outperformed it in terms of F1 score and accuracy rate.

Our method relied on two key components: wavelet reconstruction and convolutional neural networks. We showed that using wavelet reconstruction to preprocess the data, in combination with a high-pass filter and PSD, improved the clarity of the PV(OP) plot and allowed for better detection of stiction. This is because wavelet reconstruction can remove noise from the data and make the shape of the PV(OP) plot clearer. We also demonstrated that using both the filtered and wavelet reconstruction data as input to the CNN model can provide better performance, as the contrast in their shapes can provide vital insights for accurate classification.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Our study also investigated the effects of different parameters. While most tests had little effect on the model's performance, the threshold for wavelet reconstruction proved to be an interesting factor. We found that setting the threshold for wavelet reconstruction to 0.5 resulted in the clearest distinction between the two cases and the best performance. Moreover, we observed that the choice of model input significantly affected performance, as shown in Table 7, where other inputs reduced F1 scores around 0.7-0.8.

The accurate automatic detection of valve stiction in control loops holds immense significance for the industrial sector. The presence of stiction can severely impact the safety, efficiency, and overall performance of industrial processes, leading to suboptimal control, increased energy consumption, and potential equipment failures. In this context, our proposed method, which combines wavelet reconstruction and convolutional neural networks, offers a highly effective solution to tackle this critical issue.

By accurately identifying stiction in various types of control loops, our method enables proactive maintenance and targeted interventions, preventing costly downtime and optimizing process control. The implications extend beyond individual control loops, as improved stiction detection can have a cascading effect, enhancing the performance of interconnected systems and ensuring smoother operations. Moreover, the proposed method equips operators and engineers with a powerful tool for identifying and addressing stiction-related issues promptly, minimizing the risk of safety incidents and maximizing the overall efficiency of industrial processes. Through our research, we have showcased the potential of our method to significantly advance stiction detection, thereby facilitating safer, more reliable, and economically viable industrial operations.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

In conclusion, our proposed method demonstrates great potential for detecting stiction in control loops, which can have a positive impact on the safety and efficiency of industrial processes. By combining wavelet reconstruction with convolutional neural networks, we have achieved impressive performance on a diverse range of data from various industries and control loop types. Our results emphasize the importance of selecting appropriate model inputs and applying field knowledge and principles when developing models, rather than relying solely on its black-box nature. Although our study focused on a specific dataset, we believe that our proposed method can be applied to other datasets and real-world scenarios as long as the data has a similar structure. Future research could investigate the use of other neural network architectures, such as transformers or attention-based models, or the incorporation of additional features or data sources to enhance performance.

Overall, we hope that our proposed method will contribute to the development of more efficient and reliable industrial processes.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

APPENDIX A

DATA GENERATION IN LITERATURES

Due to limitations of the industries, data generation is an essential step to apply computer science techniques to chemical processes. In this appendix data generation methods of the prominent research cited in this work are shown and discussed.

Firstly, in Farenzena’s work, training data were generated by simulation as shown in a diagram in Figure 1 where the controller is a PI controller discussed in section 2.1, the valve was represented by Choudhury’s 2-parameter valve stiction model discussed in section 2.2, and the process was represented by a first order transfer function in Laplace domain as shown in equation A.1. Parameters in each component in the diagram as well as variance of white noise that adds right after the process are varied for each simulation scenario as shown in Table A.1 resulted in a dataset of 360,000 unique scenarios.

$$G_p = \frac{1}{1+\tau s} \tag{A.1}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table A.1 Variable parameters used in the Farenzena’s data generation

Parameter	Description	Interval
K_c	Controller gain	[0.1:0.2:1.1, 2:1:5]
τ_i	Controller time constant	[0.3T:0.3T,3T]
T	Process time constant	[5:5:20, 30:10:100]
S	Static band	[0.5:0.5:5]
J	Slip Jump	[0.5:0.5:5]
σ^2	White noise variance	[100 10 1]

Even though a lot of data were generated in Farenzena's work, most of those data are redundance and represent some scenarios that will not happen in the real world. So, in Amiruddin's work data are created using a single process transfer function as shown in equation A.2 with 1 second sampling time, and to represent each real-world scenario including well-tuned and tightly tuned non-stiction cases, process disturbed by external oscillations non-stiction cases and stiction cases with parameter in Table A.2, A.3 and A.4 respectively.

$$G_p = \frac{z^{-3}(1.45z - 1)}{z - 0.8} \tag{A.2}$$

Table A.2 *Variable parameters used in the Amiruddin's data generation in well-tuned and tightly tuned non-stiction cases*

Parameter	Description	Interval
K_c	Controller gain	[0.1: 0.01: 0.3]
I $(1/\tau_i)$	Final controller integral value	[0.01: 0.01: 0.27]
σ^2	White noise variance	[0, 0.010, 0.020, 0.030, 0.040, 0.050]

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table A.3 *Variable parameters used in the Amiruddin's data generation in process disturbed by external oscillations non-stiction cases with K_c and I equal to 0.15 and 0.15 s^{-1} respectively*

Parameter	Description	Interval
A	Amplitude	[1, 1.5, 2, 2.5]
f	Frequency of oscillation, in rad/s	[0.01: 0.01: 0.11]
σ^2	White noise variance	[0, 0.010, 0.020, 0.030, 0.040, 0.050]
Φ	Phase/lag of oscillation, in rad	[0, 0.25Π, 0.5Π, 0.75Π, Π, 1.25Π, 1.5Π, 1.75Π]

Table A.4 Variable parameters used in the Amiruddin's data generation in stiction cases

Parameter	Description	Interval
S	Static band	[0.1: 0.25: 10]
J	Slip Jump	[0.1: 0.25: 10]
σ^2	White noise variance	[0, 0.010, 0.020, 0.030, 0.040, 0.050]

And in Zhang's work, types of process are also considered as 2 transfer functions were used to represent self-regulated and ramp processes as shown in equation A.3 and A.4 respectively. Furthermore, controller input commands were also added in forms of sine wave and step change. Variable parameters used in this work are shown in Table A.5.

$$G_1 = \frac{1}{0.2s+1} \tag{A.3}$$

$$G_2 = \frac{1}{0.2s} e^{-0.05s} \tag{A.4}$$

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Table A.5 Variable parameters used in the Zhang’s data generation

Process	Input	K_c	I	S	J	Std_{noise}
G_1	$\sin(2\Pi t)+2$	[4: 0.2: 8]	[0.05: 0.02: 0.2]	0	0	[0.005, 0.01]
G_1	$\sin(2\Pi t)+2$	4	0.1	[0.1: 0.1: 0.7, 1.0: 0.5: 5.0]	[0.0: 0.1: 1.0]	0
G_1	$\varepsilon(t-0.05)$	[4: 0.2: 8]	[0.05: 0.02: 0.2]	0	0	[0.005, 0.01]
G_1	$\varepsilon(t-0.05)$	6	0.2	[0.1: 0.1: 1.5]	[0.00: 0.01: 0.05]	0
G_2	$\sin(2\Pi t)+2$	[4: 0.2: 8]	[0.1: 0.1: 0.5]	0	0	[0.005, 0.01]
G_2	$\sin(2\Pi t)+2$	6	0.1	[0.1: 0.1: 0.7, 1.0: 0.5: 5.0]	[0.0: 0.1: 1.0]	0
G_2	$\varepsilon(t-0.05)$	[4: 0.2: 8]	[0.05: 0.02: 0.2]	0	0	[0.005, 0.01]
G_2	$\varepsilon(t-0.05)$	6	0.5	[0.1: 0.1: 1.5]	[0.00: 0.01: 0.05]	0

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

REFERENCES

1. Raul, M. *Stiction: The hidden menace*. 2000; Available from: http://www.expertune.com/articles/RuelNov2000/stiction.htm.
2. Zheng, D., et al., *Valve Stiction Detection and Quantification Using a K-Means Clustering Based Moving Window Approach.* Industrial & Engineering Chemistry Research, 2021. **60**(6): p. 2563-2577.
3. Bialkowski, W.L. *Dreams versus reality: a view from both sides of the gap: manufacturing excellence with come only through engineering excellence*. 1993.
4. Desborough, L., P. Nordh, and R. Miller, *Control system reliability: Process out of control.* Industrial Computing, 2001. **8**: p. 52-55.
5. Paulonis, M. and J. Cox, *A Practical Approach for Large-Scale Controller Performance Assessment, Diagnosis, and Improvement.* Journal of Process Control, 2003. **13**: p. 155-168.
6. Choudhury, M.A.A.S., D.S. Shook, and S.L. Shah, *LINEAR OR NONLINEAR? A BICOHERENCE BASED METRIC OF NONLINEARITY MEASURE.* IFAC Proceedings Volumes, 2006. **39**(13): p. 617-622.
7. Xu, Z., C. Zhan, and S. Zhang, *A new non-invasive method for valve stiction dection using wavelet technology.* Journal of Electronics (China), 2009. **26**(5): p. 673-680.
8. Choudhury, M.A.A.S., N.F. Thornhill, and S.L. Shah, *A Data-Driven Model for Valve Stiction.* IFAC Proceedings Volumes, 2004. **37**(1): p. 245-250.
9. Zhu, Y. *Multivariable System Identification For Process Control*. 2001.
10. Butterworth, S., *On the Theory of Filter Amplifiers.* Experimental Wireless and the Wireless Engineer, 1930. **7**: p. 536-541.
11. Akansu, A. and Y. Liu, *On-signal decomposition techniques.* Optical Engineering, 1991. **30**(7).
12. Xu, Y., et al., *Wavelet transform domain filters: A spatially selective noise filtration technique.* Image Processing, IEEE Transactions on, 1994. **3**: p. 747-758.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

13. Kano, M., et al., *Practical Model and Detection Algorithm for Valve Stiction.* IFAC Proceedings Volumes, 2004. **37**(9): p. 859-864.

14. He, Q.P., et al., *A Curve Fitting Method for Detecting Valve Stiction in Oscillating Control Loops.* Industrial & Engineering Chemistry Research, 2007. **46**(13): p. 4549-4560.

15. Singhal, A. and T.I. Salsbury, *A simple method for detecting valve stiction in oscillating control loops.* Journal of Process Control, 2005. **15**(4): p. 371-382.

16. Rossi, M. and C. Scali, *A comparison of techniques for automatic detection of stiction: simulation and application to industrial data.* Journal of Process Control, 2005. **15**(5): p. 505-514.

17. Horch, A., *A simple method for detection of stiction in control valves.* Control Engineering Practice, 1999. **7**(10): p. 1221-1231.

18. Jelali, M., *Estimation of valve stiction in control loops using separable least-squares and global search algorithms.* Journal of Process Control, 2008. **18**(7): p. 632-642.

19. Lee, K.H., Z. Ren, and B. Huang. *Novel closed-loop stiction detection and quantification method via system identification*. in *ADCONIP Conference, Jasper, Canada*. 2008.

20. Karra, S. and M.N. Karim, *Comprehensive methodology for detection and diagnosis of oscillatory control loops.* Control Engineering Practice, 2009. **17**(8): p. 939-956.

21. Kamaruddin, B., et al., *A simple model-free butterfly shape-based detection (BSD) method integrated with deep learning CNN for valve stiction detection and quantification.* Journal of Process Control, 2020. **87**: p. 1-16.

22. Jelali, M. and B. Huang, *Detection and Diagnosis of Stiction in Control Loops: State of the Art and Advanced Methods. London: Springer-Verlag*. 2010.

23. Dambros, J.W., M. Farenzena, and J.O. Trierweiler, *Signal preprocessing for stiction detection methods.* Industrial & Engineering Chemistry Research, 2018. **57**(1): p. 302-315.

24. Farenzena, M. and J.O. Trierweiler, *A novel technique to estimate valve stiction based on pattern recognition*, in *Computer aided chemical engineering*. 2009,

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Elsevier. p. 1191-1196.

25. Zabiri, H. and M. Ramasamy, *NLPCA as a diagnostic tool for control valve stiction.* Journal of Process Control, 2009. **19**(8): p. 1368-1376.

26. Zabiri, H., et al., *NN-based algorithm for control valve stiction quantification.* WSEAS Trans. Syst. Control, 2009. **4**(2): p. 88-97.

27. Venceslau, A.R., L.A. Guedes, and D.R. Silva. *Artificial neural network approach for detection and diagnosis of valve stiction*. in *Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012)*. 2012. IEEE.

28. Mohd Amiruddin, A.A.A., et al., *Valve stiction detection through improved pattern recognition using neural networks.* Control Engineering Practice, 2019. **90**: p. 63-84.

29. Dambros, J.W.V., M. Farenzena, and J.O. Trierweiler, *Oscillation Detection and Diagnosis in Process Industries by Pattern Recognition Technique.* IFAC-PapersOnLine, 2019. **52**(1): p. 299-304.

30. Zhang, K., et al., *Multiple-Timescale Feature Learning Strategy for Valve Stiction Detection Based on Convolutional Neural Network.* IEEE/ASME Transactions on Mechatronics, 2022. **27**(3): p. 1478-1488.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

VITA

NAME	Kris Prasopsanti
DATE OF BIRTH	19 July 1990
PLACE OF BIRTH	Bangkok
INSTITUTIONS ATTENDED	Chulalongkorn University
HOME ADDRESS	239/25 Lumpini Pathumwan Bangkok 10330

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY